CURRICULUM VITAE

PERSONAL DETAILS

Name:	John Coughlin Rockett III
Nationality:	USA
Work Address:	United States Environmental Protection Agency National Health and Environmental Effects Research Laboratory Reproductive Toxicology Division (MD-72) Gamete and Early Embryo Biology Branch Research Triangle Park NC 27711 USA
Work Telephone:	+001 (919) 541 2678
Work Fax:	+001 (919) 541 4017
Work E-mail:	rockett.john@epa.gov

Employment and Higher Education

CURRENT POSITION (12/00-present)

Research Biologist
Gamete and Early Embryo Biology Branch (MD-72)
Reproductive Toxicology Division
National Health and Environmental Effects Research Laboratory
US Environmental Protection Agency
Research Triangle Park
NC 27711
USA

PREVIOUS POSITIONS

8/98-12/00: NHEERL Post-Doctoral Research Fellow, Gamete and Early Embryo Biology Branch, Reproductive Toxicology Division, National Health and Environmental Effects Research Laboratory, United States Environmental Protection Agency, Research Triangle Park, NC, USA.
Supervisors: Dr Sally P. Darney (Scientific publications under Sally D. Perreault) and Dr David J. Dix.

5/95-7/98: Rhone-Poulenc Post-Doctoral Research Fellow, Molecular Toxicology Group, School of Biological Sciences, University of Surrey, Guildford, Surrey, England.
Supervisor: Prof. G. Gordon Gibson.

EDUCATION

Ph.D., 1995 - University of Warwick, Coventry, W. Midlands, England
Title: Transforming Growth Factor-β and Immune Recognition Molecules in Oesophageal Cancer.
Supervisors: Dr Alan G. Morris (University of Warwick) and Dr S. Jane Darnton (Birmingham Heartlands Hospital)

B.Sc. (Hons.), 1991 - University of Warwick, Coventry, W. Midlands, England.
Degree: Microbiology and Microbial Technology (with intercalated year in industry), Class 2i.
Tutor: Professor Howard Dalton.

PROFESSIONAL ACTIVITIES

Membership of Professional Societies:

Society of Toxicology (Inc. Molecular Biology Speciality Section) (2001-present)
Science Advisory Board (2001-present)
North Carolina Chapter of the Society of Toxicology (1999-present)
Triangle Consortium for Reproductive Biology (1999-present)
Triangle Array Users Group (1999-present)
Institute of Biology (U.K.) (1989 - present)
British Toxicology Society (1996 - 2000)
Biochemical Society (U.K.) (1992-1995)
British Society for Immunology (1992-1995)

Membership of Scientific Committees:

International Life Sciences Institute's (ILSI) Health and Environmental Sciences Institute (HESI) Technical Committee on the Application of Genomics to Mechanism-Based Risk Assessment:

- Steering Committee (5/02-present).
- Hepatotoxicity Working Group Vice-Chair (5/02-present).
- Hepatotoxicity Work Group Member (5/01-present).

Charter member, Fertility and Early Pregnancy Work Group of the National Children's Study (07/01-Present).

National Health and Environmental Effects Research Laboratory Distinguished Lecture Series Committee (July 03-present).

U.S. Environmental Protection Agency Genomics Task Force Microarray Technical Subcommittee (August 03-present).

National Health and Environmental Effects Research Laboratory Genomics and Proteomics Committee (NGPC) (September 03-present).

Professional Meetings:

Invited participant ("Observer") in Expert Panel Workshop: "The Role of Environmental Factors on the Onset and Progression of Puberty in Children". Organised by Serono Symposia International. November 6th-8th, 2003, Chicago, IL, USA.

Joint organiser and co-chair of: "*Genomic analysis of surrogate tissues for measuring toxic exposures and drug action*", the "*Innovations in Applied Toxicology*" Symposium for the Society of Toxicology 42nd Annual Meeting, March 9th-13th, 2003, Salt Lake City, UT, USA.

(8) **John C. Rockett,** David J. Esdaile and G Gordon Gibson (1999). Differential gene expression in drug metabolism: practicalities, problems and potential. *Xenobiotica*, **29**(7):655-691.
(7) MC Murphy, CN Brookes, **JC Rockett**, C Chapman, JA Lovegrove, BJ Gould, JW Wright and CM Williams (1999). The quantitation of lipoprotein lipase mRNA in biopsies of human adipose tissue, using the polymerase chain reaction, and the effect of increased consumption of n-3 polyunsaturated fatty acids. *European Journal of Clinical Nutrition*, **53**:441-447.

(6) **JC Rockett,** DJ Esdaile and GG Gibson (1997). Molecular profiling of non-genotoxic carcinogenesis using differential display reverse transcription polymerase chain reaction (ddRT-PCR*). European Journal of Drug Metabolism & Pharmacokinetics* **22**(4):329-33.

(5) **Rockett, J.,** Larkin, K., Darnton, S., Morris, A. and Matthews, H. (1997). Five newly established oesophageal carcinoma cell lines: phenotypic and immunological characterisation. *British Journal of Cancer* **75**(2):258-263.

(4) **J C Rockett,** S J Darnton, J Crocker, H R Matthews and A G Morris (1996). Lymphocyte infiltration in oesophageal carcinoma: lack of correlation with MHC antigens, ICAM-1, and tumour stage and grade. *Journal of Clinical Pathology* **49**:264-267.

(3) **J C Rockett,** S J Darnton, J Crocker, H R Matthews and A G Morris (1995). Expression of HL-ABC and HLA-DR histocompatability antigens and intercellular adhesion molecule-1 in oesophageal carcinoma. *Journal of Clinical Pathology* **48**:539-44.

(2) Salam M, **Rockett J** and Morris A (1995). The prevalence of different human papillomavirus types and p53 mutations in laryngeal carcinomas: is there a reciprocal relationship? *European Journal of Surgical Oncology* **21**:290-296.

(1) Salam M, **Rockett J** and Morris A (1995). General primer-mediated polymerase chain reaction for simultaneous detection and typing of HPV in laryngeal carcinomas. *Clinical Otolaryngology* **20**:84-88.

(2) Articles Submitted To A Scientific Journal

(4) **John C. Rockett**, Judith E. Schmid, Christopher J. Luft, J. Brian Garges, M. Stacey Ricci, Pasquale Patrizio, Norman B. Hecht and David J. Dix. Gene Expression Patterns Associated with Infertility in Rodent and Human Models. ***An invited submission***

(3) Roger Ulrich, **John C. Rockett**, G. Gordon Gibson and Syril Pettit. Evaluating the Effects of Methapyrilene and Clofibrate on Hepatic Gene Expression: A Collaboration Between Laboratories and a Comparison of Platform and Analytical Approaches.

(2) Valerie A Baker, Helen M Harries, Jeffrey F Waring, Roger Jolly, Angus de Souza, Judith E Schmid, Hong Ni, Roger Brown, Roger G Ulrich and **John C. Rockett**. Clofibrate-Induced Gene Expression Changes in Rat Liver: A Cross-Laboratory Analysis Using Membrane cDNA Arrays.

(1) David Miller, Corrado Spadafora, David Dix, Adrian Platts, **John C. Rockett**, Stephen A Krawetz Nuclease digestion of sperm chromatin suggests a random distribution of gene sequences.

(3) Articles In Preparation For Submission To A Scientific Journal

(3) Spearow J, DB Tully, **JC Rockett** and DJ Dix. Differential testicular gene expression in mouse strains sensitive and resistant to endocrine disruption by estrogen.

(2) Sally D. Perrault, **John C. Rockett**, Laura Fenster, James Kesner, Wendy Robbins and Steven Schrader. Biomarkers for Assessing Reproductive Development and Health: Part 2 – Adult Reproductive Health.

(1) J. Christopher Luft, Douglas B. Tully, **John C. Rockett**, Judith E. Schmid and David J. Dix. Reproductive and genomic effects in testes from mice exposed to the water disinfectant byproduct bromochloroacetic acid

(4) Book Chapters

(4) **John C. Rockett**. Gene Microarrays Applied to Reproductive Toxicology. In Cunningham (Ed): *Genetic and Proteomic Applications in Toxicity Testing*, The Human Press, Totowa. In Preparation. ****An invited submission****

(3) **John C. Rockett** and David J Dix. Gene Expression Networks. In Cooper (ed-in-chief): *Encyclopaedia of the Human Genome*, Nature Publishing Group. London, New York. ISBN 0-333-80386-8 (2003). ****An invited submission****

(2) **John C. Rockett**. The Future of Toxicogenomics. In Michael Burczynski (ed): "*An Introduction to Toxicogenomics*". CRC Press. Boca Raton, London, New York, Washington D.C., pp299-317 (2003). ****An invited submission****

(1) **J. Rockett**, S. Darnton, J. Crocker, H. Matthews and A. Morris: Major Histocompatibility Complex (MHC) class I and II and Intercellular Adhesion Molecule (ICAM)-1 expression in oesophageal carcinoma. Peracchia A, Rosati R, Bonavina L, Bona S, Chella B (eds): *Recent Advances in Diseases of the Esophagus*. Bologna: Monduzzi Editore, pp45-49 (1996).

(5) Other Scientific Publications (Letters to Editors; Meeting Reports; Commentaries etc.)

(11) **John C. Rockett** (2003). Probing the nature of microarray-based oligonucleotides. *Drug Discovery Today* 8(9):389. *(A Letter To The Editor)* ****An invited submission****

(10) **John C. Rockett** (2003). To confirm or not to confirm (microarray data) – that, is the question. *Drug Discovery Today* 8(8):343. *(A Letter To The Editor)*

(9B) Nazzareno Ballatori, James L. Boyer, and **John C. Rockett**. (2003). Exploiting Genome Data to Understand the Function, Regulation and Evolutionary Origins of Toxicologically Relevant Genes. *Environ Health Perspect*. 111(6):871-5. *(A Meeting Report)*

(9A) Nazzareno Ballatori, James L. Boyer, and **John C. Rockett**. (2003). Exploiting Genome Data to Understand the Function, Regulation and Evolutionary Origins of Toxicologically Relevant Genes. *EHP Toxicogenomics*. 111(1T):61-5. *(A Meeting Report)*

(8) **John C. Rockett** (2002). Surrogate Tissue Analysis for Monitoring the Degree and Impact of Exposures in Agricultural Workers. *AgBiotechNet,* 4:1-7 November, ABN 100. *(A Review Article)*. ***An invited submission***

(7) **John C. Rockett** (2002). Macroresults Through Microarrays. *Drug Discovery Today*, 7(15);804-805. *(A Meeting Report)*

(6) **John C. Rockett** (2002). Chip, chip, array! Three chips for post-genomic research. *Drug Discovery Today*, 7(8);458-459. *(A Meeting Report)*

(5) **John C. Rockett** (2002). Use of Genomic Data in Risk Assessment. *Genome***Biology**, 3(4): reports4011.1-4011.3 (http://genomebiology.com/2002/3/4/reports/4011/?isguard=1). *(A Meeting Report)*

(4) **John C. Rockett** (2001). Genomic and Proteomic Techniques Applied to Reproductive Biology. *Genome***Biology** 2(9): 4020.1-4020.3 (http://genomebiology.com/2001/2/9/reports/4020/). *(A Meeting Report)*

(3) **John C. Rockett** (2001). Chipping away at the mystery of drug responses. *The Pharmacogenomics Journal*, 1(3);161-163. *(A commentary)* ***An invited submission***

(2) **Rockett, John C**. and Dix, David J. (1999). U.S. EPA workshop: Application of DNA arrays to Toxicology. *Environmental Health Perspectives*, **107**(8):681-685. *(A Meeting Report)*

(1) **John C. Rockett III** (1995). Immune recognition molecules and transforming growth factor beta-1 in oesophageal cancer. Ph.D. thesis, University of Warwick, Coventry, England.*(Ph.D. thesis)*

(6) Published Book, Paper and Website reviews

(9) **John C. Rockett** (2002). A report on the manuscript: Systemic RNAi in *C. elegans* requires the putative transmembrane protein SID-1. Winston WM, Molodowitch C, Hunter CP. *Science.* 2002 **295**:2456-2459. *Genome***Biology**, **3**(7):reports0034 http://genomebiology.com/2002/3/7/reports/0034/

(8) **John C. Rockett** (2001). A report on the manuscript: Genetic rescue of an endangered mammal by cross-species nuclear transfer using post-mortem somatic cells. P Loi , et al., *Nat Biotechnol.* 2001, 19:962-964. *Genome***Biology**, **3**(1):reports0006. (http://genomebiology.com/2001/3/1/reports/0006/).

(7) **John C. Rockett** (2001). A report on the manuscript: Molecular Classification of Human Carcinomas by Use of Gene Expression Signatures. A Su et al., *Cancer Res.* 2001 61:7388-7393. *Genome***Biology**, **3**(1):reports0005. (http://genomebiology.com/2001/3/1/reports/0005/).

(6) **John C. Rockett** (2001). A report on the manuscript: Genetic evidence for two species of elephant in Africa. A Roca et al., Science. 2001 Aug 24;293(5534):1473-7. *Genome***Biology**, **2**(12):reports0045. (http://www.genomebiology.com/2001/2/12/reports/0045/.

(5) **John C. Rockett** (2001). A report on the manuscript: Extensive genetic polymorphism in the human CYP2B6 gene with impact on expression and function in human liver. T Lang et al., *Pharmacogenetics*, 2001, 11(5):399-415. *Genome***Biology**, **2**(12):reports0044. (http://www.genomebiology.com/2001/2/12/reports/0044/).

(4) **John C. Rockett** (2001). A report on the manuscript: Novel Human Testis-Specific cDNA: molecular Cloning, Expression and Immunological Effects of the Recombinant Protein. R Santhanam and R K Naz, *Molecular Reproduction and Development* 60:1-12 (2001). *Genome***Biology**, **2**(11):reports0040. (http://genomebiology.com/2001/2/11/reports/0040/).

(3) **John C. Rockett** (2001). A report on the website: BIND - The Biomolecular Interaction Network Database (http://www.bind.ca/). *Genome***Biology**, **2**(9): reports2011. http://www.genomebiology.com/2001/2/9/reports/2011/.

(2) **John C. Rockett** (2001). A report on the manuscript: Exploring the DNA-binding specificities of zinc fingers with DNA microarrays. ML Bulyk et al., *Proc Natl Acad Sci USA* 2001, **98**:7158-7163. *Genome***Biology**, **2**(10): reports0032. (http://genomebiology.com/2001/2/10/reports/0032/).

(1) J **Rockett** (1996). A Book Review on: "Cell Adhesion and Cancer" (Eds., Hogg N. and Hart I.). Clinical Molecular Pathology 49(1):M64. ****An invited submission****

(7) Published Abstracts of Poster and Oral Presentations

(17) Amber K. Goetz, Wenjun Bao, Judith E. Schmid, Carmen Wood, Hongzu Ren, Deborah S. Best, Rachel N. Murrell, **John C. Rockett**, Michael G. Narotsky, Douglas C. Wolf, Douglas B. Tully, David J. Dix Gene Expression Profiling in Testis and Liver of Mice to Identify Modes of Action of Conazole Toxicities. Society of Toxicology 43rd Annual Meeting, March 21st-25th, 2004, Baltimore, MD, USA. *Toxicological Sciences*. (Submitted)

(16) Jane Gallagher, Theresa Lehman, Ramakrishna Modali, Scott Rhoney, Marien Clas, Jeff Inmon, **John C. Rockett**, David Dix, Cindy Mamay, Suzanne Fenton, Suzanne McMaster, Stan

Barone Jr, Pauline Mendola and Reeder Sams. Validation of Non-Invasive Biological Samples: Pilot Projects Relevant to the National Children Study. Society of Toxicology 43rd Annual Meeting, March 21st-25th, 2004, Baltimore, MD, USA. *Toxicological Sciences*. (Submitted)

(15) B.S. Pukazhenthi, **J. C. Rockett**, M. Ouyang, D.J. Dix, J.G. Howard, P. Georgopoulos, W.J. J. Welsh and D. E. Wildt. Gene Expression In The Testis Of Normospermic Versus Teratospermic Domestic Cats Using Human cDNA Microarray Analyses. Society for the Study of Reproduction 36th Annual Meeting, July 19th-22nd, 2003, Cincinnati, OH, USA. *Biology of Reproduction* **68** (Supp 1):191.

(14) David J. Dix and **John C. Rockett** (2003). Genomic and Proteomic Analysis of Surrogate Tissues for Assessing Toxic Exposures and Disease States. Innovation in Applied Toxicology symposium entitled "*Genomic and Proteomic Analysis of Surrogate Tissues for Assessing Toxic Exposures and Disease States*". Society of Toxicology 42nd Annual Meeting, March 9th-13th, 2003, Salt Lake City, UT, USA. *Toxicological Sciences* 72(S-1):276.

(13) **John C. Rockett**, Chad R. Blystone, Amber K. Goetz, Rachel N. Murrell, Judith E. Schmid and David J. Dix. (2003). Gene Expression Profiling Of Accessible Surrogate Tissues To Monitor Molecular Changes In Inaccessible Target Tissues Following Toxicant Exposure. Innovations in Applied Toxicology Symposium entitled "*Genomic and Proteomic Analysis of Surrogate Tissues for Assessing Toxic Exposures and Disease States*". Society of Toxicology 42nd Annual Meeting, March 9th-13th, 2003, Salt Lake City, UT, USA. *Toxicological Sciences* 72(S-1):276.

(12) Douglas B. Tully, J. Christopher Luft, **John C. Rockett**, Judy E. Schmid and David J. Dix (2002). Effects on gene expression in testes from adult male mice exposed to the water disinfectant byproduct bromochloroacetic acid. *Society for the Study of Reproduction 35th Annual Meeting, July 28-31, 2002, Baltimore, Maryland, USA. Biology of Reproduction* 66 (Supp 1):223.

(11) David J. Dix, Kary E. Thompson, **John C. Rockett**, Judith E. Schmid, Robert J. Goodrich, David Miller, G. Charles Ostermeier and Stephen A. Krawetz (2002). Testis and spermatazola RNA profiles of normal fertile men. *Society for the Study of Reproduction 35th Annual Meeting, July 28-31, 2002, Baltimore, Maryland, USA. Biology of Reproduction* 66 (Supp 1):194.

(10) Asa J. Oudes, **John C. Rockett**, David J. Dix and Kwan Hee Kim (2002). Identification of retinoic acid induced genes in mouse testis by cDNA microarray analysis. *27th Annual Meeting of the American Society of Andrology*, 4/24-27/02. *J. Andrology* Supplement March/April.

(9) **John C. Rockett**, Robert J. Kavlock, Christy Lambright, Louise G. Parks, Judith E. Schmid, Vickie S. Wilson and David J. Dix (2002). Use of DNA arrays to monitor gene expression in blood and uterus from Long-Evans rats following 17-β-estradiol exposure – a new approach to biomonitoring endocrine disrupting chemicals using surrogate tissues. *Toxicological Sciences* 66(1): Abstract No.1388.

(8) David J. Dix and **John C. Rockett** (2002). Genomic analysis of the testicular toxicity of haloacetic acids. Platform presentation at the symposium, "Defining the cellular and molecular

mechanisms of toxicant action in the testis". *Toxicological Science* 66 (1): Abstract No.848.

(7) **JC Rockett,** JC Luft, JB Garges and DJ Dix (2001). The reproductive effects of the water disinfectant byproduct bromochloroacetate on juvenile and adult male mice. *Toxicological Sciences*, **60** (1):250.

(6) Tarka DK, Klinefelter GR, **Rockett JC**, Suarez JD, Roberts NL and Rogers JM (2001). Effect of gestational expsore to ethane dimethane sulfonate (EDS), bromochloroacetic acid (BCA) and molinate on reproductive function in CD-1 male mice. *Toxicological Sciences*, 60 (1):250.

(5) Garges JB, **Rockett JC** and Dix DJ (2001). Developmental and reproductive phenotype of mice lacking stress-inducible 70 kDa heat shock proteins (Hsp70s). *Toxicological Sciences*, 60 (1):383.
(4) D Dix, **J Rockett,** J Luft, J Garges, M Ricci, P Patrizio and N Hecht (2000). Using DNA microarrays to characterise gene expression in testes of fertile and infertile humans and mice. *Biology of Reproduction*, 62 (s1);227.

(3) J Luft, J B Garges, **J Rockett** and D Dix (2000). Male reproductive toxicity of bromochloroacetic acid in mice. *Biology of Reproduction*, 62 (s1);246.

(2) **Rockett, JC**, Garges, JB and Dix, DJ (2000). A single heat-shock of juvenile male mice causes a long-term decrease in fertility and reduces embryo quality. *Toxicological Sciences* 54 (1):365.

(1) **JC Rockett**, SJ Darnton, J Crocker, HR Matthews and AG Morris (1994). Major Histocompatability (MHC) class I and II and intercellular adhesion molecule (ICAM)-1 expression in oesophageal carcinoma (OC). *Immunology* 83 (s1):64.

(8) Invited Oral Presentations

(10) **John C. Rockett** and Gary M Hellmann. *To confirm or not to confirm (microarray data) – that is the question.* Seminar for EPA/NHEERL Genomics and Proteomics Committee's ArrayQA forum, August 25th, 2003, RTP, NC, USA.

(9) **John C. Rockett**. "*Biomonitoring Toxicant Exposure and Effect Using Toxicogenomics and Surrogate Tissue Analysis*". Seminar for Division of Epidemiology, Statistics and Prevention Research, National Institute of Child Health and Development, May 29th, 2003, Rockville, MD, USA.

(8) **John C. Rockett**. "*Genomics and Proetomics: New Toxicity Testing*". Platform presentation at US EPA Regional Risk Assessors Annual Conference, April 28th – May 2nd, 2003, Stone Mountain, GA, USA.

(7) **John C. Rockett,** Chad R. Blystone, Amber K. Goetz, Rachel N. Murrell, Judith E. Schmid and David J. Dix. "*Gene Expression Profiling Of Accessible Surrogate Tissues To Monitor Molecular Changes in Inaccessible Target Tissues Following Toxicant Exposure.*" Platform presentation at

SoT 42nd Annual Meeting symposium entitled "*Genomic and Proteomic Analysis of Surrogate Tissues for Measuring Toxic Exposures and Drug Action*", March 9th-13th, 2003, Salt Lake City, UT, USA.

(6) **John C. Rockett**. "*A Toxicogenomic Approach to Surrogate Tissue Analysis*". Seminar for Department of Environmental and Molecular Toxicology, North Carolina State University, September 3rd, 2002, Raleigh, NC, USA.

(5) **John C. Rockett**. "Differential gene expression in toxicology: practicalities, problems and potential". Platform presentation at *9th Annual Mount Desert Island Biological Laboratory Environmental Health Sciences Symposium: Exploiting Genome Data to Understand the Function, Regulation and Evolutionary Origins of Toxicologically Relevant Genes*, July 10th-11th, 2002, Salisbury Cove, Maine, USA.

(4) **John C. Rockett**, Leroy Folmar, Michael J. Hemmer and David J. Dix. "Arrays for biomonitoring environmental and reproductive toxicology". Platform Presentation at *Macroresults Through Microarrays 3 – Advancing Drug Development*, April 29th-May 1st, 2002, Boston, MA, USA.

(3) **John C. Rockett**, Sigmund Degitz, Suzanne E. Fenton, Leroy Folmar, Michael J. Hemmer, Joe E Tietge, and David J. Dix. "Use of DNA Arrays in Environmental Toxicology". Platform presentation at the *4th Annual Lab-on-a-Chip and Microarrays for Post-Genomic Applications meeting*, January 14th-16th, 2002, Zurich, Switzerland.

(2) **John C. Rockett**. "DNA Arrays". Seminar at *EPA Molecular Biology Course*, April 8th, 1999, USEPA, RTP, NC, USA.

(1) **John C. Rockett.** "Contract Services for Array Applications". Seminar at the *Triangle Array Users Group*, May 1st, 1999, CIIT, RTP, NC, USA.

(9) Other Poster and Oral Presentations

(23) **John C. Rockett**, Wenjun Bao, Chad R. Blystone, Amber K. Goetz, Rachel N. Murrell, Hongzu Ren, Judith E. Schmid, Jessica Stapelfeldt, Lillian F. Strader, Kary E. Thompson and David J. Dix. Genomic Analysis of Surrogate Tissues for Assessing Environmental Exposures and Future Disease States. ILSI-HESI meeting: *Toxicogenomics in Risk Assessment - Assessing the Utility, Challenges, and Next Steps*. June 5th-6th, 2003, Fairfax, VA, USA.

(22) **John C. Rockett**, Wenjun Bao, Chad R. Blystone, Amber K. Goetz , Rachel N. Murrell, Hongzu Ren, Judith E. Schmid, Jessica Stapelfeldt, Lillian F. Strader, Kary E. Thompson and David J. Dix. Genomic Analysis of Surrogate Tissues for Assessing Environmental Exposures and Future Disease States. *EPA Science Forum*, May 5th-7th, 2003, Washington, D.C., USA.

(21) Germaine Buck, Courtney Johnson, Joseph Stanford, Anne Sweeney, Laura Schieve, **John Rockett**, Sherry Selevan and Steve Schrader. Prospective Pregnancy Study Designs for Assessing Reproductive and Developmental Toxicants. *American Epidemiology Society Meeting*, March 27th-28th, 2003, Atlanta, GA, USA.

(20) **John C. Rockett**, Chad R. Blystone, Amber K. Goetz, Rachel N. Murrell, Hongzu Ren, Judith E. Schmid, Jessica Stapelfeldt, Lillian F. Strader, Kary E. Thompson, Doug B. Tully, Paul Zigas and David J. Dix. Genomic Analysis of Surrogate Tissues for Assessing Environmental Exposures and Future Disease States. *National Children's Study Assembly Meeting*, December 16th-18th, 2002, Baltimore, MD, USA.

(19) **John Rockett**. The Use of Gene Expression Profiling to Detect Early Biomarkers of Adverse Effects Prior to Clinical manifestation. *National Children's Study: Meeting of EPA Project Leaders* – Methods Development Projects. November 20th, 2002, USEPA, RTP, NC, USA. (Oral Presentation)

(18) GC Ostermeier, RJ Goodrich, K Thompson, **J Rockett**, MP Diamond, K Collins, NICHD Reproductive Medicine Network, DJ. Dix, D Miller and SA Krawetz. Defining the spermatozoal RNA population in normal fertile men. *American Society of Reproductive Medicine* October 12-17, 2002, Seattle, WA, USA.

(17) G. Charles Ostermeier, Robert J. Goodrich, Kary Thompson, **John Rockett,** Michael P. Diamond, Karen Collins, NICHD Reproductive Medicine Network, David J. Dix, David Miller and Stephen A. Krawetz. RNAs isolated from ejaculate spermatozoa provide a noninvasive means to investigate testicular gene expression. *Gordon Conference on Mammalian Gametogenesis & Embryogenesis*, June 30th-July 5th, Connecticut College, New London, CT, USA.

(16) David Dix, **John Rockett**, Judith Schmid, Lillian Strader, Douglas Tully. Genomic analysis of testicular toxicity. *USEPA/NHEERL/RTD Peer Review*, October 22nd, 2001, RTP, NC, USA.

(15) David Dix, **John Rockett**, Judith Schmid, Douglas Tully. Monitoring human reproductive health and development through gene expression profiling. *USEPA/NHEERL/RTD Peer Review*, October 22nd, 2001, RTP, NC, USA.

(14) Patrizio P, N Hecht, **J Rockett**, J Schmid and D Dix (2001). DNA microarrays to study gene expression profiles in testis of fertile and infertile men. *57th Annual Meeting of the American Society for Reproductive Medicine*, October 20th-25th, 2001, Orlando, FL, USA.

(13) Jimmy L. Spearow, Dale Morris, Uland Wong, Rashid Altafi, Saeed Eteiwi, Mark Stanford, Trevor Stearns, Lorena Orozio, Angela Chen, **John Rockett**, Douglas Tully, David Dix and Marylynn Barkley. Genetic Variation In Susceptibility To The Disruption Of Testicular Development And Gene Expression By Pubertal Exposure To Estrogenic Agents. *Third Annual University of California at Davis Conference for Environmental Health Scientists, Disruption of Developing Systems and Advances in Therapeutic Approaches* August 27th, 2001, UC Davis, CA, USA.

(12) Tarka DK, Klinefelter GR, **Rockett JC**, Suarez JD, Roberts NL and Rogers JM (2001). Effect of gestational expsore to ethane dimethane sulfonate (EDS), bromochloroacetic acid (BCA) and molinate on reproductive function in CD-1 male mice. *North Carolina Society of Toxicology Winter Meeting*, March 3rd, 2001. NIEHS, RTP, NC, USA.

(11) David Dix, **John Rockett**, Leroy Folmar, Michael Hemmer, Sigmund Degitz, and Joseph Tietge (2001). Biomonitoring the Toxicogenomic Response to Endocrine Disrupting Chemicals in Humans, Laboratory Species and Wildlife. *U.S. - Japan International Workshop for Endocrine Disrupting Chemicals*, February 28th-March 3rd, 2001, Tsukuba, Japan.

(10) **John C. Rockett**, Faye L. Mapp, J. Brian Garges, J. Christopher Luft, Chisato Mori and David J Dix David Dix (2001). The effects of hyperthermia on spermatogenesis, apoptosis, gene expression and fertility in adult male mice. *Triangle Consortium for Reproductive Biology Annual Meeting*, January 27th, 2001, RTP, NC, USA.

(9) Gangolli E, Dix DJ, Garges J B, **Rockett, JC** and Idzerda RL (2000). Testosterone Regulation of Sertoli Cell genes. *11th International Congress of Endocrinology*, October 29th-November 2nd, 2000, Sydney, Australia.

(8) **J Rockett**, J Luft, J Garges, M Ricci, P Patrizio, N Hecht and D Dix (2000). Using DNA microarrays to characterise gene expression in testes of fertile and infertile humans and mice. *Functional Genomics & Microarray Data Mining*, August 3rd-4th 2000, Durham, NC, USA.

(7) **Rockett JC**, S Ricci, P Patrizio, NB Hecht, JB Garges and DJ Dix (2000). Gene Expression in the Mammalian Testis. *5th NHEERL Symposium*, June 6th-8th, 2000, RTP, NC, USA.

(6) J Luft, J B Garges, **J Rockett** and D Dix (2000). Male reproductive toxicity of bromochloroacetic acid in mice. *2000 NIEHS/NTA Biomedical Science and Career Fair*, April 28th 2000, RTP, NC, USA.

(5) **Rockett JC**, S Ricci, P Patrizio, NB Hecht, JB Garges and DJ Dix (2000). Gene Expression in the Mammalian Testis. *Molecular Toxicology, Toxicogenomics and Associated Bioinformatics Applied to Drug Discovery* meeting, January 11th-15th, 2000,Santa Fe,NM, USA.

(4) **JC Rockett** and DJ Dix (1999). Development of DNA arrays for the analysis of testis-expressed genes in humans and mice. *The 8th Annual National Health and Environmental Effects Research Laboratory Open House*. November 2nd-3rd, 1999, RTP, NC, USA.

(3) **JC Rockett**, DJ Esdaile and GG Gibson (1997). Molecular profiling of non-genotoxic carcinogenesis using differential display reverse transcription polymerase chain reaction (ddRT-PCR). *The British Toxicology Society Annual Meeting*, April 19th-22nd, 1998, University of Surrey, Guildford, Surrey, England.

(2) **JC Rockett**, DJ Esdaile and GG Gibson (1997). Molecular profiling of non-genotoxic

carcinogenesis using differential display reverse transcription polymerase chain reaction (ddRT-PCR). Poster presentation at *Symposium on Drug Metabolism: Towards the next Millennium*. August 26th-28th,1997, London King's College, London, England.

(1) **J Rockett**, S Darnton, J Crocker, H Matthews and A Morris: Major Histocompatibility Complex (MHC) class I and II and Intercellular Adhesion Molecule (ICAM)-1 expression in oesophageal carcinoma. Oral presentation at *The 6th World Congress of the International Society for Diseases of the Esophagus*, August 23rd-26th, 1995, Milan, Italy.

Axl1p sequence following Ser[209] and occurs within the domain of Axl1p that shows homology with hIDE (14). To delete the complete *STE23* sequence and create the *ste23Δ::URA3* mutation, polymerase chain reaction (PCR) primers (5′-TCGGAAGACCTCATTCTTGCTCATTTTGATATTGCTC- TGTAGATTGTACTGAGAGTGCAC-3′; and 5′-GCTACAAACAGCGTCGACTTGAATGCCCCGACATCTTCGACTGTGCGGTATTTCACACCG-3′) were used to amplify the *URA3* sequence of pRS316, and the reaction product was transformed into yeast for one-step gene replacement [R. Rothstein, *Methods Enzymol.* 194, 281 (1991)]. To create the *axl1Δ::LEU2* mutation contained on p114, a 5.0-kb Sal I fragment from p*AXL1* was cloned into pUC19, and an internal 4.0-kb Hpa I–Xho I fragment was replaced with a *LEU2* fragment. To construct the *ste23Δ::LEU2* allele (a deletion corresponding to 931 amino acids) carried on p153, a *LEU2* fragment was used to replace the 2.8-kb Pml I–Ecl136 II fragment of *STE23*, which occurs within a 6.2-kb Hind III–Bgl II genomic fragment carried on pSP72 (Promega). To create YEp*MFA1*, a 1.6-kb Bam HI fragment containing *MFA1*, from pKK16 [K. Kuchler, R. E. Sterne, J. Thorner, *EMBO J.* 8, 3973 (1989)], was ligated into the Bam HI site of YEp351 [J. E. Hill, A. M. Myers, T. J. Koerner, A. Tzagoloff, *Yeast* 2, 163 (1986)].

24. J. Chant and I. Herskowitz, *Cell* 65, 1203 (1991).
25. B. W. Matthews, *Acc. Chem. Res.* 21, 333 (1988).
26. K. Kuchler, H. G. Dohlman, J. Thorner, *J. Cell Biol.* 120, 1203 (1993); R. Kolling and C. P. Hollenberg, *EMBO J.* 13, 3261 (1994); C. Berkower, D. Loayza, S. Michaelis, *Mol. Biol. Cell* 5, 1185 (1994).
27. A. Bender and J. R. Pringle, *Proc. Natl. Acad. Sci. U.S.A.* 86, 9976 (1989); J. Chant, K. Corrado, J. R. Pringle, I. Herskowitz, *Cell* 65, 1213 (1991); S. Powers, E. Gonzales, T. Christensen, J. Cubert, D. Broek, *ibid.*, p. 1225; H. O. Park, J. Chant, I. Herskowitz, *Nature* 365, 269 (1993); J. Chant, *Trends Genet.* 10, 328 (1994); ______ and J. R. Pringle, *J. Cell Biol.* 129, 751 (1995); J. Chant, M. Mischke, E. Mitchell, I. Herskowitz, J. R. Pringle, *ibid.*, p. 767.
28. G. F. Sprague Jr., *Methods. Enzymol.* 194, 77 (1991).
29. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
30. A W303 1A derivative, SY2625 (*MAT***a** *ura3-1 leu2-3, 112 trp1-1 ade2-1 can1-100 sst1Δ mfa2Δ::FUS1-lacZ his3Δ::FUS1-HIS3*), was the parent strain for the mutant search. SY2625 derivatives for the mating assays, secreted pheromone assays, and the pulse-chase experiments included the following strains: Y49 (*ste22-1*), Y115 (*mfa1Δ::LEU2*), Y142 (*axl1::URA3*), Y173 (*axl1Δ::LEU2*), Y220 (*axl1::URA3 ste23Δ::URA3*), Y221 (*ste23Δ::URA3*), Y231 (*axl1Δ::LEU2 ste23Δ::LEU2*), and Y233 (*ste23Δ::LEU2*). *MAT*α derivatives of SY2625 included the following strains: Y199 (SY2625 made *MAT*α), Y278 (*ste22-1*), Y195 (*mfa1Δ::LEU2*), Y196 (*axl1Δ::LEU2*), and Y197 (*axl1::URA3*). The EG123 (*MAT***a** *leu2 ura3 trp1 can1 his4*) genetic background was used to create a set of strains for analysis of bud site selection. EG123 derivatives included the following strains: Y175 (*axl1Δ::LEU2*), Y223 (*axl1::URA3*), Y234 (*ste23Δ::LEU2*), and Y272 (*axl1Δ::LEU2 ste23Δ::LEU2*). *MAT*α derivatives of EG123 included the following strains: Y214 (EG123 made *MAT*α) and Y293 (*axl1Δ::LEU2*). All strains were generated by means of standard genetic or molecular methods involving the appropriate constructs (23). In particular, the *axl1 ste23* double mutant strains were created by crossing of the appropriate *MAT***a** *ste23* and *MAT*α *axl1* mutants, followed by sporulation of the resultant diploid and isolation of the double mutant from nonparental di-type tetrads. Gene disruptions were confirmed with either PCR or Southern (DNA) analysis.
31. p129 is a YEp352 [J. E. Hill, A. M. Myers, T. J. Koerner, A. Tzagoloff, *Yeast* 2, 163 (1986)] plasmid containing a 5.5-kb Sal I fragment of p*AXL1*. p151 was derived from p129 by insertion of a linker at the Bgl II site within *AXL1*, which led to an in-frame insertion of the hemagglutinin (HA) epitope (DQYPYDVPDYA) (29) between amino acids 854 and 855 of the *AXL1* product. pC225 is a KS+ (Stratagene) plasmid containing a 0.5-kb Bam HI–Sst I fragment from p*AXL1*. Substitution mutations of the proposed active site of Axl1p were created with the use of pC225 and site-specific mutagenesis involving appropriate synthetic oligonucleotides (*axl1-H68A*, 5′-GTGCTCACAAAGCGCTGCCAAACCGGC-3′; *axl1-E71A*, 5′-AAGAATCATGTGCGCACAAAGGTGCGC-3′; and *axl1-E71D*, 5′-AAGAATCATGTGATCACAAAGGTGCGC-3′). The mutations were confirmed by sequence analysis. After mutagenesis, the 0.4-kb Bam HI–Msc I fragment from the mutagenized pC225 plasmids was transferred into p*AXL1* to create a set of pRS316 plasmids carrying different *AXL1* alleles, p124 (*axl1-H68A*), p130 (*axl1-E71A*), and p132 (*axl1-E71D*). Similarly, a set of HA-tagged alleles carried on YEp352 were created after replacement of the p151 Bam HI–Msc I fragment, to generate p161 (*axl1-E71A*), p162 (*axl1-*

32.

N. Davis, T. Favero, C. de Hoog, and S. Kim for comments on the manuscript. Supported by a grant to C.B. from the Natural Sciences and Engineering Research Council of Canada. Support for M.N.A. was from a California Tobacco-Related Disease Research Program postdoctoral fellowship (4FT-0083).

22 June 1995; accepted 21 August 1995

Quantitative Monitoring of Gene Expression Patterns with a Complementary DNA Microarray

Mark Schena,* Dari Shalon,*† Ronald W. Davis, Patrick O. Brown‡

A high-capacity system was developed to monitor the expression of many genes in parallel. Microarrays prepared by high-speed robotic printing of complementary DNAs on glass were used for quantitative expression measurements of the corresponding genes. Because of the small format and high density of the arrays, hybridization volumes of 2 microliters could be used that enabled detection of rare transcripts in probe mixtures derived from 2 micrograms of total cellular messenger RNA. Differential expression measurements of 45 *Arabidopsis* genes were made by means of simultaneous, two-color fluorescence hybridization.

The temporal, developmental, topographical, histological, and physiological patterns in which a gene is expressed provide clues to its biological role. The large and expanding database of complementary DNA (cDNA) sequences from many organisms (*1*) presents the opportunity of defining these patterns at the level of the whole genome.

For these studies, we used the small flowering plant *Arabidopsis thaliana* as a model organism. *Arabidopsis* possesses many advantages for gene expression analysis, including the fact that it has the smallest genome of any higher eukaryote examined to date (*2*). Forty-five cloned *Arabidopsis* cDNAs (Table 1), including 14 complete sequences and 31 expressed sequence tags (ESTs), were used as gene-specific targets. We obtained the ESTs by selecting cDNA clones at random from an *Arabidopsis* cDNA library. Sequence analysis revealed that 28 of the 31 ESTs matched sequences in the database (Table 1). Three additional cDNAs from other organisms served as controls in the experiments.

The 48 cDNAs, averaging ~1.0 kb, were amplified with the polymerase chain reaction (PCR) and deposited into individual wells of a 96-well microtiter plate. Each sample was duplicated in two adjacent wells to allow the reproducibility of the arraying and hybridization process to be tested. Samples from the microtiter plate were printed onto glass microscope slides in an area measuring 3.5 mm by 5.5 mm with the use of a high-speed arraying machine (*3*). The arrays were processed by chemical and heat treatment to attach the DNA sequences to the glass surface and denature them (*3*). Three arrays, printed in a single lot, were used for the experiments here. A single microtiter plate of PCR products provides sufficient material to print at least 500 arrays.

Fluorescent probes were prepared from total *Arabidopsis* mRNA (*4*) by a single round of reverse transcription (*5*). The *Arabidopsis* mRNA was supplemented with human acetylcholine receptor (AChR) mRNA at a dilution of 1:10,000 (w/w) before cDNA synthesis, to provide an internal standard for calibration (*5*). The resulting fluorescently labeled cDNA mixture was hybridized to an array at high stringency (*6*) and scanned

M. Schena and R. W. Davis, Department of Biochemistry, Beckman Center, Stanford University Medical Center, Stanford, CA 94305, USA.
D. Shalon and P. O. Brown, Department of Biochemistry and Howard Hughes Medical Institute, Beckman Center, Stanford University Medical Center, Stanford, CA 94305, USA.

*These authors contributed equally to this work.
†Present address: Synteni, Palo Alto, CA 94303, USA.
‡To whom correspondence should be addressed. E-mail: pbrown@cmgm.stanford.edu

with a laser (3). A high-sensitivity scan gave signals that saturated the detector at nearly all of the *Arabidopsis* target sites (Fig. 1A). Calibration relative to the AChR mRNA standard (Fig. 1A) established a sensitivity limit of ~1:50,000. No detectable hybridization was observed to either the rat glucocorticoid receptor (Fig. 1A) or the yeast *TRP4* (Fig. 1A) targets even at the highest scanning sensitivity. A moderate-sensitivity scan of the same array allowed linear detection of the more abundant transcripts (Fig. 1B). Quantitation of both scans revealed a range of expression levels spanning three orders of magnitude for the 45 genes tested (Table 2). RNA blots (7) for several genes (Fig. 2) corroborated the expression levels measured with the microarray to within a factor of 5 (Table 2).

Differential gene expression was investigated with a simultaneous, two-color hybridization scheme, which served to minimize experimental variation inherent in the comparison of independent hybridizations. Fluorescent probes were prepared from two mRNA sources with the use of reverse transcriptase in the presence of fluorescein- and lissamine-labeled nucleotide analogs, respectively (5). The two probes were then mixed together in equal proportions, hybridized to a single array, and scanned separately for fluorescein and lissamine emission after independent excitation of the two fluorophores (3).

To test whether overexpression of a single gene could be detected in a pool of total *Arabidopsis* mRNA, we used a microarray to analyze a transgenic line overexpressing the single transcription factor *HAT4* (8). Fluorescent probes representing mRNA from wild-type and *HAT4*-transgenic plants were labeled with fluorescein and lissamine, respectively; the two probes were then mixed and hybridized to a single array. An intense hybridization signal was observed at the position of the *HAT4* cDNA in the lissamine-specific scan (Fig. 1D), but not in the fluorescein-specific scan of the same array (Fig. 1C). Calibration with AChR mRNA added to the fluorescein and lissamine cDNA synthesis reactions at dilutions of 1:10,000 (Fig. 1C) and 1:100 (Fig. 1D), respectively, revealed a 50-fold elevation of *HAT4* mRNA in the transgenic line relative to its abundance in wild-type plants (Table 2). This magnitude of *HAT4* overexpression matched that inferred from the Northern (RNA) analysis within a factor of 2 (Fig. 2 and Table 2). Expression of all the other genes monitored on the array differed by less than a factor of 5 between *HAT4*-transgenic and wild-type plants (Fig 1, C

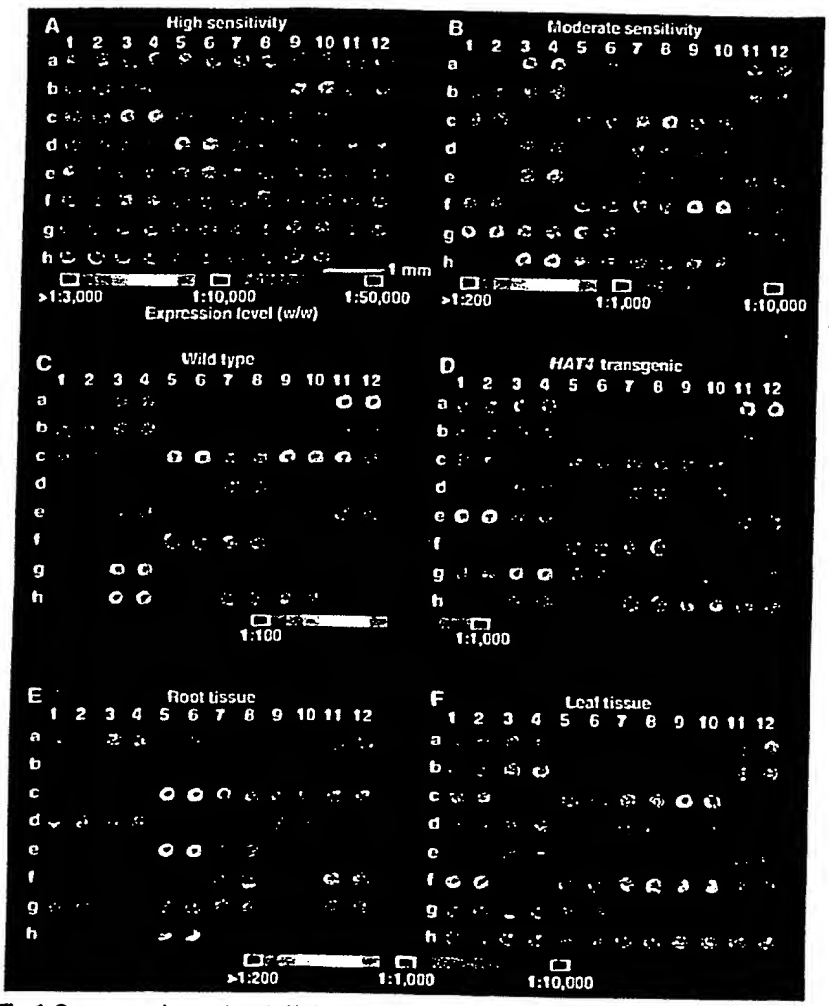


Fig. 1. Gene expression monitored with the use of cDNA microarrays. Fluorescent scans represented in pseudocolor correspond to hybridization intensities. Color bars were calibrated from the signal obtained with the use of known concentrations of human AChR mRNA in independent experiments. Numbers and letters on the axes mark the position of each cDNA. **(A)** High-sensitivity fluorescein scan after hybridization with fluorescein-labeled cDNA derived from wild-type plants. **(B)** Same array as in (A) but scanned at moderate sensitivity. **(C and D)** A single array was probed with a 1:1 mixture of fluorescein-labeled cDNA from wild-type plants and lissamine-labeled cDNA from HAT4-transgenic plants. The single array was then scanned successively to detect the fluorescein fluorescence corresponding to mRNA from wild-type plants (C) and the lissamine fluorescence corresponding to mRNA from HAT4-transgenic plants (D). **(E and F)** A single array was probed with a 1:1 mixture of fluorescein-labeled cDNA from root tissue and lissamine-labeled cDNA from leaf tissue. The single array was then scanned successively to detect the fluorescein fluorescence corresponding to mRNAs expressed in roots (E) and the lissamine fluorescence corresponding to mRNAs expressed in leaves (F).

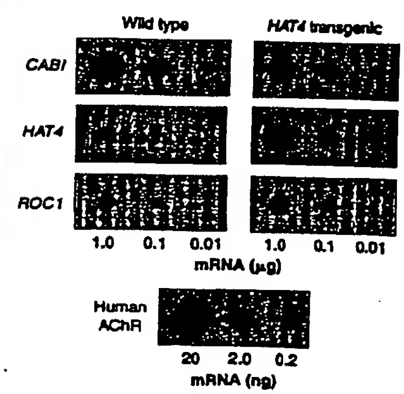


Fig. 2. Gene expression monitored with RNA (Northern) blot analysis. Designated amounts of mRNA from wild-type and *HAT4*-transgenic plants were spotted onto nylon membranes and probed with the cDNAs indicated. Purified human AChR mRNA was used for calibration.

and D, and Table 2). Hybridization of fluorescein-labeled glucocorticoid receptor cDNA (Fig. 1C) and lissamine-labeled *TRP4* cDNA (Fig. 1D) verified the presence of the negative control targets and the lack of optical cross talk between the two fluorophores.

To explore a more complex alteration in expression patterns, we performed a second two-color hybridization experiment with fluorescein- and lissamine-labeled probes prepared from root and leaf mRNA, respectively. The scanning sensitivities for the two fluorophores were normalized by matching the signals resulting from AChR mRNA, which was added to both cDNA synthesis reactions at a dilution of 1:1000 (Fig. 1, E and F). A comparison of the scans revealed widespread differences in gene expression between root and leaf tissue (Fig. 1, E and F). The mRNA from the light-regulated *CABI* gene was ~500-fold more abundant in leaf (Fig. 1F) than in root tissue (Fig. 1E). The expression of 26 other genes differed between root and leaf tissue by more than a factor of 5 (Fig. 1, E and F).

The *HAT4*-transgenic line we examined has elongated hypocotyls, early flowering, poor germination, and altered pigmentation (*8*). Although changes in expression were observed for *HAT4*, large changes in expression were not observed for any of the other 44 genes we examined. This was somewhat surprising, particularly because comparative analysis of leaf and root tissue identified 27 differentially expressed genes. Analysis of an expanded set of genes may be required to identify genes whose expression changes upon *HAT4* overexpression; alternatively, a comparison of mRNA populations from specific tissues of wild-type and *HAT4*-transgenic plants may allow identification of downstream genes.

At the current density of robotic printing, it is feasible to scale up the fabrication process to produce arrays containing 20,000 cDNA targets. At this density, a single array would be sufficient to provide gene-specific targets encompassing nearly the entire repertoire of expressed genes in the *Arabidopsis* genome (*2*). The availability of 20,274 ESTs from *Arabidopsis* (*1*, *9*) would provide a rich source of templates for such studies.

The estimated 100,000 genes in the human genome (*10*) exceeds the number of *Arabidopsis* genes by a factor of 5 (*2*). This modest increase in complexity suggests that similar cDNA microarrays, prepared from the rapidly growing repertoire of human ESTs (*1*), could be used to determine the expression patterns of tens of thousands of human genes in diverse cell types. Coupling an amplification strategy to the reverse transcription reaction (*11*) could make it feasible to monitor expression even in minute tissue samples. A wide variety of acute and chronic physiological and pathological conditions might lead to characteristic changes in the patterns of gene expression in peripheral blood cells or other easily sampled tissues. In concert with cDNA microarrays for monitoring complex expression patterns, these tissues might therefore serve as sensitive in vivo sensors for clinical diagnosis. Microarrays of cDNAs could thus provide a useful link between human gene sequences and clinical medicine.

Table 1. Sequences contained on the cDNA microarray. Shown is the position, the known or putative function, and the accession number of each cDNA in the microarray (Fig. 1). All but three of the ESTs used in this study matched a sequence in the database. NADH, reduced form of nicotinamide adenine dinucleotide; ATPase, adenosine triphosphatase; GTP, guanosine triphosphate.

Position	cDNA	Function	Accession number
a1, 2	AChR	Human AChR	*
a3, 4	EST3	Actin	H36236
a5, 6	EST6	NADH dehydrogenase	Z27010
a7, 8	*AAC1*	Actin 1	M20016
a9, 10	EST12	Unknown	U36594†
a11, 12	EST13	Actin	T45783
b1, 2	*CABI*	Chlorophyll a/b binding	M85150
b3, 4	EST17	Phosphoglycerate kinase	T44490
b5, 6	*GA4*	Gibberellic acid biosynthesis	L37126
b7, 8	EST19	Unknown	U36595†
b9, 10	*GBF-1*	G-box binding factor 1	X63894
b11, 12	EST23	Elongation factor	X52256
c1, 2	EST29	Aldolase	T04477
c3, 4	*GBF-2*	G-box binding factor 2	X63895
c5, 6	EST34	Chloroplast protease	R87034
c7, 8	EST35	Unknown	T14152
c9, 10	EST41	Catalase	T22720
c11, 12	rGR	Rat glucocorticoid receptor	M14053
d1, 2	EST42	Unknown	U36596†
d3, 4	EST45	ATPase	J04185
d5, 6	*HAT1*	Homeobox-leucine zipper 1	U09332
d7, 8	EST46	Light harvesting complex	T04063
d9, 10	EST49	Unknown	T76267
d11, 12	*HAT2*	Homeobox-leucine zipper 2	U09335
e1, 2	*HAT4*	Homeobox-leucine zipper 4	M90394
e3, 4	EST50	Phosphoribulokinase	T04344
e5, 6	*HAT5*	Homeobox-leucine zipper 5	M90416
e7, 8	EST51	Unknown	Z33675
e9, 10	*HAT22*	Homeobox-leucine zipper 22	U09336
e11, 12	EST52	Oxygen evolving	T21749
f1, 2	EST59	Unknown	Z34607
f3, 4	*KNAT1*	*Knotted*-like homeobox 1	U14174
f5, 6	EST60	RuBisCO small subunit	X14564
f7, 8	EST69	Translation elongation factor	T42799
f9, 10	*PPH1*	Protein phosphatase 1	U34803
f11, 12	EST70	Unknown	T44621
g1, 2	EST75	Chloroplast protease	T43698
g3, 4	EST78	Unknown	R65481
g5, 6	*ROC1*	Cyclophilin	L14844
g7, 8	EST82	GTP binding	X59152
g9, 10	EST83	Unknown	Z33795
g11, 12	EST84	Unknown	T45278
h1, 2	EST91	Unknown	T13832
h3, 4	EST96	Unknown	R64816
h5, 6	*SAR1*	Synaptobrevin	M90418
h7, 8	EST100	Light harvesting complex	Z18205
h9, 10	EST103	Light harvesting complex	X03909
h11, 12	*TRP4*	Yeast tryptophan biosynthesis	X04273

*Proprietary sequence of Stratagene (La Jolla, California). †No match in the database; novel EST.

Table 2. Gene expression monitoring by microarray and RNA blot analyses; tg, *HAT4*-transgenic. See Table 1 for additional gene information. Expression levels (w/w) were calibrated with the use of known amounts of human AChR mRNA. Values for the microarray were determined from microarray scans (Fig. 1); values for the RNA blot were determined from RNA blots (Fig. 2).

Gene	Expression level (w/w)	
	Microarray	RNA blot
CABI	1:48	1:83
CABI (tg)	1:120	1:150
HAT4	1:8300	1:6300
HAT4 (tg)	1:150	1:210
ROC1	1:1200	1:1800
ROC1 (tg)	1:260	1:1300

REFERENCES AND NOTES

1. The current EST database (dbEST release 091495) from the National Center for Biotechnology Information (Bethesda, MD) contains a total of 322,225 entries, including 255,645 from the human genome and 21,044 from *Arabidopsis*. Access is available via the World Wide Web (http://www.ncbi.nlm.nih.gov).
2. E. M. Meyerowitz and R. E. Pruitt, *Science* 229, 1214 (1985); R. E. Pruitt and E. M. Meyerowitz, *J. Mol. Biol.* 187, 169 (1986); I. Hwang *et al.*, *Plant J.* 1, 367 (1991); P. Jarvis *et al.*, *Plant Mol. Biol.* 24, 685 (1994); L. Le Guen *et al.*, *Mol. Gen. Genet.* 245, 390 (1994).
3. D. Shalon, thesis, Stanford University (1995); ——— and P. O. Brown, in preparation. Microarrays were fabricated on poly-L-lysine–coated microscope slides (Sigma) with a custom-built arraying machine fitted with one printing tip. The tip loaded 1 μl of PCR product (0.5 mg/ml) from 96-well microtiter plates and deposited ~0.005 μl per slide on 40 slides at a spacing of 500 μm. The printed slides were rehydrated for 2 hours in a humid chamber, snap-dried at 100°C for 1 min, rinsed in 0.1% SDS, and treated with 0.05% succinic anhydride prepared in buffer consisting of 50% 1-methyl-2-pyrrolidinone and 50% boric acid. The cDNA on the slides was denatured in distilled water for 2 min at 90°C immediately before use. Microarrays were scanned with a laser fluorescent scanner that contained a computer-controlled XY stage and a microscope objective. A mixed gas, multiline laser allowed sequential excitation of the two fluorophores. Emitted light was split according to wavelength and detected with two photomultiplier tubes. Signals were read into a PC with the use of a 12-bit analog-to-digital board. Additional details of microarray fabrication and use may be obtained by means of e-mail (pbrown@cmgm. stanford.edu).
4. F. M. Ausubel *et al.*, Eds., *Current Protocols in Molecular Biology* (Greene & Wiley Interscience, New York, 1994), pp. 4.3.1–4.3.4.
5. Polyadenylated [poly(A)$^+$] mRNA was prepared from total RNA with the use of Oligotex-dT resin (Qiagen). Reverse transcription (RT) reactions were carried out with a StrataScript RT-PCR kit (Stratagene) modified as follows: 50-μl reactions contained 0.1 μg/μl of *Arabidopsis* mRNA, 0.1 ng/μl of human AChR mRNA, 0.05 μg/μl of oligo(dT) (21-mer), 1× first strand buffer, 0.03 U/μl of ribonuclease block, 500 μM deoxyadenosine triphosphate (dATP), 500 μM deoxyguanosine triphosphate, 500 μM dTTP, 40 μM deoxycytosine triphosphate (dCTP), 40 μM fluorescein-12-dCTP (or lissamine-5-dCTP), and 0.03 U/μl of StrataScript reverse transcriptase. Reactions were incubated for 60 min at 37°C, precipitated with ethanol, and resuspended in 10 μl of TE (10 mM tris-HCl and 1 mM EDTA, pH 8.0). Samples were then heated for 3 min at 94°C and chilled on ice. The RNA was degraded by adding 0.25 μl of 10 N NaOH followed by a 10-min incubation at 37°C. The samples were neutralized by addition of 2.5 μl of 1 M tris-Cl (pH 8.0) and 0.25 μl of 10 N HCl and precipitated with ethanol. Pellets were washed with 70% ethanol, dried to completion in a speedvac, resuspended in 10 μl of H_2O, and reduced to 3.0 μl in a speedvac. Fluorescent nucleotide analogs were obtained from New England Nuclear (DuPont).
6. Hybridization reactions contained 1.0 μl of fluorescent cDNA synthesis product (5) and 1.0 μl of hybridization buffer [10× saline sodium citrate (SSC) and 0.2% SDS]. The 2.0-μl probe mixtures were aliquoted onto the microarray surface and covered with cover slips (12 mm round). Arrays were transferred to a hybridization chamber (3) and incubated for 18 hours at 65°C. Arrays were washed for 5 min at room temperature (25°C) in low-stringency wash buffer (1× SSC and 0.1% SDS), then for 10 min at room temperature in high-stringency wash buffer (0.1× SSC and 0.1% SDS). Arrays were scanned in 0.1× SSC with the use of a fluorescence laser-scanning device (3).
7. Samples of poly(A)$^+$ mRNA (4, 5) were spotted onto nylon membranes (Nytran) and crosslinked with ultraviolet light with the use of a Stratalinker 1800 (Stratagene). Probes were prepared by random priming with the use of a Prime-It II kit (Stratagene) in the presence of [^{32}P]dATP. Hybridizations were carried out according to the instructions of the manufacturer. Quantitation was performed on a PhosphorImager (Molecular Dynamics).
8. M. Schena and R. W. Davis, *Proc. Natl. Acad. Sci. U.S.A.* 89, 3894 (1992); M. Schena, A. M. Lloyd, R. W. Davis, *Genes Dev.* 7, 367 (1993); M. Schena and R. W. Davis, *Proc. Natl. Acad. Sci. U.S.A.* 91, 8393 (1994).
9. H. Hofte *et al.*, *Plant J.* 4, 1051 (1993); T. Newman *et al.*, *Plant Physiol.* 106, 1241 (1994).
10. N. E. Morton, *Proc. Natl. Acad. Sci. U.S.A.* 88, 7474 (1991); E. D. Green and R. H. Waterston, *J. Am. Med. Assoc.* 266, 1966 (1991); C. Bellanne-Chantelot, *Cell* 70, 1059 (1992); D. R. Cox *et al.*, *Science* 265, 2031 (1994).
11. E. S. Kawasaki *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 85, 5698 (1988).

12. The laser fluorescent scanner was designed and fabricated in collaboration with S. Smith of Stanford University. Scanner and analysis software was developed by R. X. Xia. The succinic anhydride reaction was suggested by J. Mulligan and J. Van Ness of Darwin Molecular Corporation. Thanks to S. Theologis, C. Somerville, K. Yamamoto, and members of the laboratories of R.W.D. and P.O.B. for critical comments. Supported by the Howard Hughes Medical Institute and by grants from NIH [R21HG00450 (P.O.B.) and R37AG00198 (R.W.D.)] and from NSF (MCB9106011) (R.W.D.) and by an NSF graduate fellowship (D.S.). P.O.B. is an assistant investigator of the Howard Hughes Medical Institute.

11 August 1995; accepted 22 September 1995

Gene Therapy in Peripheral Blood Lymphocytes and Bone Marrow for ADA$^-$ Immunodeficient Patients

Claudio Bordignon,* Luigi D. Notarangelo, Nadia Nobili, Giuliana Ferrari, Giulia Casorati, Paola Panina, Evelina Mazzolari, Daniela Maggioni, Claudia Rossi, Paolo Servida, Alberto G. Ugazio, Fulvio Mavilio

Adenosine deaminase (ADA) deficiency results in severe combined immunodeficiency, the first genetic disorder treated by gene therapy. Two different retroviral vectors were used to transfer ex vivo the human ADA minigene into bone marrow cells and peripheral blood lymphocytes from two patients undergoing exogenous enzyme replacement therapy. After 2 years of treatment, long-term survival of T and B lymphocytes, marrow cells, and granulocytes expressing the transferred ADA gene was demonstrated and resulted in normalization of the immune repertoire and restoration of cellular and humoral immunity. After discontinuation of treatment, T lymphocytes, derived from transduced peripheral blood lymphocytes, were progressively replaced by marrow-derived T cells in both patients. These results indicate successful gene transfer into long-lasting progenitor cells, producing a functional multilineage progeny.

Severe combined immunodeficiency associated with inherited deficiency of ADA (*1*) is usually fatal unless affected children are kept in protective isolation or the immune system is reconstituted by bone marrow transplantation from a human leukocyte antigen (HLA)–identical sibling donor (*2*). This is the therapy of choice, although it is available only for a minority of patients. In recent years, other forms of therapy have been developed, including transplants from haploidentical donors (*3, 4*), exogenous enzyme replacement (*5*), and somatic-cell gene therapy (*6–9*).

We previously reported a preclinical model in which ADA gene transfer and expression successfully restored immune functions in human ADA-deficient (ADA$^-$) peripheral blood lymphocytes (PBLs) in immunodeficient mice in vivo (*10, 11*). On the basis of these preclinical results, the clinical application of gene therapy for the treatment of ADA$^-$ SCID (severe combined immunodeficiency disease) patients who previously failed exogenous enzyme replacement therapy was approved by our Institutional Ethical Committees and by the Italian National Committee for Bioethics (*12*). In addition to evaluating the safety and efficacy of the gene therapy procedure, the aim of the study was to define the relative role of PBLs and hematopoietic stem cells in the long-term reconstitution of immune functions after retroviral vector–mediated ADA gene transfer. For this purpose, two structurally identical vectors expressing the human ADA complementary DNA (cDNA), distinguishable by the presence of alternative restriction sites in a nonfunctional region of the viral long-terminal repeat (LTR), were used to transduce PBLs and bone marrow (BM) cells independently. This procedure allowed identification of the origin of

C. Bordignon, N. Nobili, G. Ferrari, D. Maggioni, C. Rossi, P. Servida, F. Mavilio, Telethon Gene Therapy Program for Genetic Diseases, DIBIT, Istituto Scientifico H. S. Raffaele, Milan, Italy.
L. D. Notarangelo, E. Mazzolari, A. G. Ugazio, Department of Pediatrics, University of Brescia Medical School, Brescia, Italy.
G. Casorati, Unità di Immunochimica, DIBIT, Istituto Scientifico H. S. Raffaele, Milan, Italy.
P. Panina, Roche Milano Ricerche, Milan, Italy.

*To whom correspondence should be addressed.

XENOBIOTICA, 1999, VOL. 29, NO. 7, 655–691

Differential gene expression in drug metabolism and toxicology: practicalities, problems and potential

JOHN C. ROCKETT†, DAVID J. ESDAILE‡ and G. GORDON GIBSON*

Molecular Toxicology Laboratory, School of Biological Sciences, University of Surrey, Guildford, Surrey, GU2 5XH, UK

Received January 8, 1999

1. An important feature of the work of many molecular biologists is identifying which genes are switched on and off in a cell under different environmental conditions or subsequent to xenobiotic challenge. Such information has many uses, including the deciphering of molecular pathways and facilitating the development of new experimental and diagnostic procedures. However, the student of gene hunting should be forgiven for perhaps becoming confused by the mountain of information available as there appears to be almost as many methods of discovering differentially expressed genes as there are research groups using the technique.

2. The aim of this review was to clarify the main methods of differential gene expression analysis and the mechanistic principles underlying them. Also included is a discussion on some of the practical aspects of using this technique. Emphasis is placed on the so-called 'open' systems, which require no prior knowledge of the genes contained within the study model. Whilst these will eventually be replaced by 'closed' systems in the study of human, mouse and other commonly studied laboratory animals, they will remain a powerful tool for those examining less fashionable models.

3. The use of suppression-PCR subtractive hybridization is exemplified in the identification of up- and down-regulated genes in rat liver following exposure to phenobarbital, a well-known inducer of the drug metabolizing enzymes.

4. Differential gene display provides a coherent platform for building libraries and microchip arrays of 'gene fingerprints' characteristic of known enzyme inducers and xenobiotic toxicants, which may be interrogated subsequently for the identification and characterization of xenobiotics of unknown biological properties.

Introduction

It is now apparent that the development of almost all cancers and many non-neoplastic diseases are accompanied by altered gene expression in the affected cells compared to their normal state (Hunter 1991, Wynford-Thomas 1991, Vogelstein and Kinzler 1993, Semenza 1994, Cassidy 1995, Kleinjan and Van Hegningen 1998). Such changes also occur in response to external stimuli such as pathogenic micro-organisms (Rohn *et al.* 1996, Singh *et al.* 1997, Griffin and Krishna 1998, Lunney 1998) and xenobiotics (Sewall *et al.* 1995, Dogra *et al.* 1998, Ramana and Kohli 1998), as well as during the development of undifferentiated cells (Hecht 1998, Rudin and Thompson 1998, Schneider-Maunoury *et al.* 1998). The potential medical and therapeutic benefits of understanding the molecular changes which occur in any given cell in progressing from the normal to the 'altered' state are enormous. Such profiling essentially provides a 'fingerprint' of each step of a

* Author for correspondence; e-mail: g.gibson@surrey.ac.uk

† Current Address: US Environmental Protection Agency, National Health and Environmental Effects, Research Laboratory, Reproductive Toxicology Division, Research Triangle Park, NC 27711, USA.

‡ Rhone-Poulenc Agrochemicals, Toxicology Department, Sophia-Antipolis, Nice, France.

Xenobiotica ISSN 0049–8254 print/ISSN 1366–5928 online © 1999 Taylor & Francis Ltd
http://www.tandf.co.uk/JNLS/xen.htm http://www.taylorandfrancis.com/JNLS/xen.htm

cell's development or response and should help in the elucidation of specific and sensitive biomarkers representing, for example, different types of cancer or previous exposure to certain classes of chemicals that are enzyme inducers.

In drug metabolism, many of the xenobiotic-metabolizing enzymes (including the well-characterized isoforms of cytochrome P450) are inducible by drugs and chemicals in man (Pelkonen *et al.* 1998), predominantly involving transcriptional activation of not only the cognate cytochrome P450 genes, but additional cellular proteins which may be crucial to the phenomenon of induction. Accordingly, the development of methodology to identify and assess the full complement of genes that are either up- or down-regulated by inducers are crucial in the development of knowledge to understand the precise molecular mechanisms of enzyme induction and how this relates to drug action. Similarly, in the field of chemical-induced toxicity, it is now becoming increasingly obvious that most adverse reactions to drugs and chemicals are the result of multiple gene regulation, some of which are causal and some of which are casually-related to the toxicological phenomenon *per se*. This observation has led to an upsurge in interest in gene-profiling technologies which differentiate between the control and toxin-treated gene pools in target tissues and is, therefore, of value in rationalizing the molecular mechanisms of xenobiotic-induced toxicity. Knowledge of toxin-dependent gene regulation in target tissues is not solely an academic pursuit as much interest has been generated in the pharmaceutical industry to harness this technology in the early identification of toxic drug candidates, thereby shortening the developmental process and contributing substantially to the safety assessment of new drugs. For example, if the gene profile in response to say a testicular toxin that has been well-characterized *in vivo* could be determined in the testis, then this profile would be representative of all new drug candidates which act via this specific molecular mechanism of toxicity, thereby providing a useful and coherent approach to the early detection of such toxicants. Whereas it would be informative to know the identity and functionality of all genes up/down regulated by such toxicants, this would appear a longer term goal, as the majority of human genes have not yet been sequenced, far less their functionality determined. However, the current use of gene profiling yields a *pattern* of gene changes for a xenobiotic of unknown toxicity which may be matched to that of well-characterized toxins, thus alerting the toxicologist to possible *in vivo* similarities between the unknown and the standard, thereby providing a platform for more extensive toxicological examination. Such approaches are beginning to gain momentum, in that several biotechnology companies are commercially producing 'gene chips' or 'gene arrays' that may be interrogated for toxicity assessment of xenobiotics. These chips consist of hundreds/thousands of genes, some of which are degenerate in the sense that not all of the genes are mechanistically-related to any one toxicological phenomenon. Whereas these chips are useful in broad-spectrum screening, they are maturing at a substantial rate, in that gene arrays are now becoming more specific, e.g. chips for the identification of changes in growth factor families that contribute to the aetiology and development of chemically-induced neoplasias.

Although documenting and explaining these genetic changes presents a formidable obstacle to understanding the different mechanisms of development and disease progression, the technology is now available to begin attempting this difficult challenge. Indeed, several 'differential expression analysis' methods have been developed which facilitate the identification of gene products that demonstrate

altered expression in cells of one population compared to another. These methods have been used to identify differential gene expression in many situations, including invading pathogenic microbes (Zhao *et al.* 1998), in cells responding to extracellular and intracellular microbial invasion (Duguid and Dinauer 1990, Ragno *et al.* 1997, Maldarelli *et al.* 1998), in chemically treated cells (Syed *et al.* 1997, Rockett *et al.* 1999), neoplastic cells (Liang *et al.* 1992, Chang and Terzaghi-Howe 1998), activated cells (Gurskaya *et al.* 1996, Wan *et al.* 1996), differentiated cells (Hara *et al.* 1991, Guimaraes *et al.* 1995a, b), and different cell types (Davis *et al.* 1984, Hedrick *et al.* 1984, Xhu *et al.* 1998). Although differential expression analysis technologies are applicable to a broad range of models, perhaps their most important advantage is that, in most cases, absolutely no prior knowledge of the specific genes which are up- or down-regulated is required.

The field of differential expression analysis is a large and complex one, with many techniques available to the potential user. These can be categorized into several methodological approaches, including:

(1) Differential screening,
(2) Subtractive hybridization (SH) (includes methods such as chemical cross-linking subtraction—CCLS, suppression-PCR subtractive hybridization—SSH, and representational difference analysis—RDA),
(3) Differential display (DD),
(4) Restriction endonuclease facilitated analysis (including serial analysis of gene expression—SAGE—and gene expression fingerprinting—GEF),
(5) Gene expression arrays, and
(6) Expressed sequence tag (EST) analysis.

The above approaches have been used successfully to isolate differentially expressed genes in different model systems. However, each method has its own subtle (and sometimes not so subtle) characteristics which incur various advantages and disadvantages. Accordingly, it is the purpose of this review to clarify the mechanistic principles underlying the main differential expression methods and to highlight some of the broader considerations and implications of this very powerful and increasingly popular technique. Specifically, we will concentrate on the so-called 'open' systems, namely those which do not require any knowledge of gene sequences and, therefore, are useful for isolating unknown genes. Two 'closed' systems (those utilising previously identified gene sequences), EST analysis and the use of DNA arrays, will also be considered briefly for completeness. Whilst emphasis will often be placed on suppression PCR subtractive hybridization (SSH, the approach employed in this laboratory), it is the aim of the authors to highlight, wherever possible, those areas of common interest to those who use, or intend to use, differential gene expression analysis.

Differential cDNA library screening (DS)

Despite the development of multiple technological advances which have recently brought the field of gene expression profiling to the forefront of molecular analysis, recognition of the importance of differential gene expression and characterization of differentially expressed genes has existed for many years. One of the original approaches used to identify such genes was described 20 years ago by St John and Davis (1979). These authors developed a method, termed 'differential plaque filter

hybridization', which was used to isolate galactose-inducible DNA sequences from yeast. The theory is simple: a genomic DNA library is prepared from normal, unstimulated cells of the test organism/tissue and multiple filter replicas are prepared. These replica blots are probed with radioactively (or otherwise) labelled complex cDNA probes prepared from the control and test cell mRNA populations. Those mRNAs which are differentially expressed in the treated cell population will show a positive signal only on the filter probed with cDNA from the treated cells. Furthermore, labelled cDNA from different test conditions can be used to probe multiple blots, thereby enabling the identification of mRNAs which are only up-regulated under certain conditions. For example, St John and Davis (1979) screened replica filters with acetate-, glucose- and galactose-derived probes in order to obtain genes induced specifically by galactose metabolism. Although groundbreaking in its time this method is now considered insensitive and time-consuming, as up to 2 months are required to complete the identification of genes which are differentially expressed in the test population. In addition, there is no convenient way to check that the procedure has worked until the whole process has been completed.

Subtractive Hybridization (SH)

The developing concept of differential gene expression and the success of early approaches such as that described by St John and Davis (1979) soon gave rise to a search for more convenient methods of analysis. One of the first to be developed was SH, numerous variations of which have since been reported (see below). In general, this approach involves hybridization of mRNA/cDNA from one population (tester) to excess mRNA/cDNA from another (driver), followed by separation of the unhybridized tester fraction (differentially expressed) from the hybridized common sequences. This step has been achieved physically, chemically and through the use of selective polymerase chain reaction (PCR) techniques.

Physical separation

Original subtractive hybridization technology involved the physical separation of hybridized common species from unique single stranded species. Several methods of achieving this have been described, including hydroxyapatite chromatography (Sargent and Dawid 1983), avidin-biotin technology (Duguid and Dinauer 1990) and oligodT-latex separation (Hara *et al.* 1991). In the first approach, common mRNA species are removed by cDNA (from test cells)-mRNA (from control cells) subtractive hybridization followed by hydroxyapatite chromatography, as hydroxy-apatite specifically adsorbs the cDNA-mRNA hybrids. The unabsorbed cDNA is then used either for the construction of a cDNA library of differentially expressed genes (Sargent and Dawid 1983, Schneider *et al.* 1988) or directly as a probe to screen a preselected library (Zimmerman *et al.* 1980, Davis *et al.* 1984, Hedrick *et al.* 1984). A schematic diagram of the procedure is shown in figure 1.

Less rigorous physical separation procedures coupled with sensitivity enhancing PCR steps were later developed as a means to overcome some of the problems encountered with the hydroxyapatite procedure. For example, Daguid and Dinauer (1990) described a method of subtraction utilizing biotin-affinity systems as a means to remove hybridized common sequences. In this process, both the control and tester mRNA populations are first converted to cDNA and an adaptor ('oligovector',

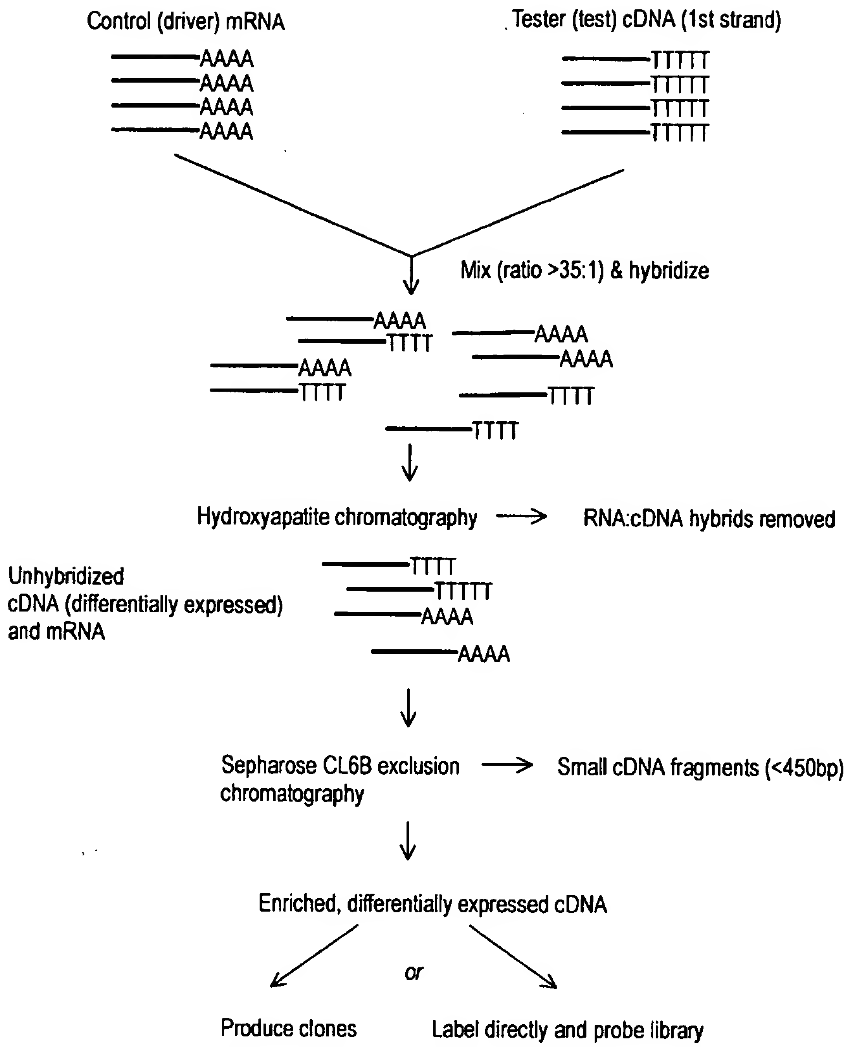


Figure 1. The hydroxyapatite method of subtractive hybridization. cDNA derived from the treated/altered (tester) population is mixed with a large excess of mRNA from the control (driver) population. Following hybridization, mRNA-cDNA hybrids are removed by hydroxyapatite chromatography. The only cDNAs which remain are those which are differentially expressed in the treated/altered population. In order to facilitate the recovery of full length clones, small cDNA fragments are removed by exclusion chromatography. The remaining cDNAs are then cloned into a vector for sequencing, or labelled and used directly to probe a library, as described by Sargent and Dawid (1983).

containing a restriction site) ligated to both sides. Both populations are then amplified by PCR, but the driver cDNA population is subsequently digested with the adaptor-containing restriction endonuclease. This serves to cleave the oligo-vector and reduce the amplification potential of the control population. The digested control population is then biotinylated and an excess mixed with tester cDNA. Following denaturation and hybridization, the mix is applied to a biocytin column (streptavidin may also be used) to remove the control population, including heteroduplexes formed by annealing of common sequences from the tester population. The procedure is repeated several times following the addition of fresh

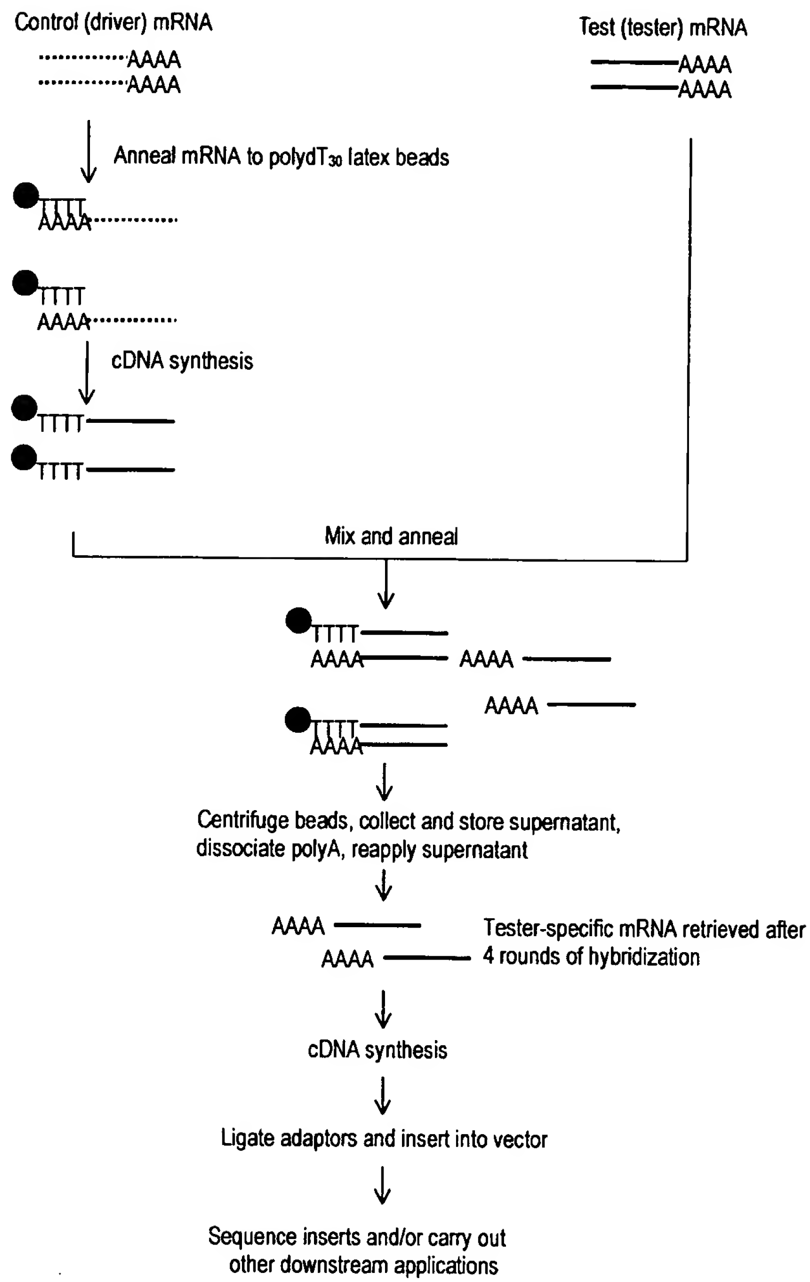


Figure 2. The use of oligodT_{30} latex to perform subtractive hybridization. mRNA extracted from the control (driver) population is converted to anchored cDNA using polydT oligonucleotides attached to latex beads. mRNA from the treated/altered (tester) population is repeatedly hybridized against an excess of the anchored driver cDNA. The final population of mRNA is tester specific and can be converted into cDNA for cloning and other downstream applications, as described by Hara *et al.* (1991).

control cDNA. In order to further enrich those species differentially expressed in the tester cDNA, the subtracted tester population is amplified by PCR following every second subtraction cycle. After six cycles of subtraction (three reamplification steps) the reaction mix is ligated into a vector for further analysis.

In a slightly different approach, Hara *et al.* (1991) utilized a method whereby oligo(dT_{30}) primers attached to a latex substrate are used to first capture mRNA extracted from the control population. Following 1st strand cDNA synthesis, the RNA strand of the heteroduplexes is removed by heat denaturation and centrifugation (the cDNA-oligotex-dT_{30} forms a pellet and the supernatant is removed). A quantity of tester mRNA is then repeatedly hybridized to the immobilized control (driver) cDNA (which is present in 20-fold excess). After several rounds of hybridization the only mRNA molecules left in the tester mRNA population are those which are not found in the driver cDNA-oligotex-dT_{30} population. These tester-specific mRNA species are then converted to cDNA and, following the addition of adaptor sequences, amplified by PCR. The PCR products are then ligated into a vector for further analysis using restriction sites incorporated into the PCR primers. A schematic illustration of this subtraction process is shown in figure 2.

However, all these methods utilising physical separation have been described as inefficient due to the requirement for large starting amounts of mRNA, significant loss of material during the separation process and a need for several rounds of hybridization. Hence, new methods of differential expression analysis have recently been designed to eliminate these problems.

Chemical Cross-Linking Subtraction (CCLS)

In this technique, originally described by Hampson *et al.* (1992), driver mRNA is mixed with tester cDNA (1st strand only) in a ratio of $>20{:}1$. The common sequences form cDNA:mRNA hybrids, leaving the tester specific species as single stranded cDNA. Instead of physically separating these hybrids, they are inactivated chemically using 2,5 diaziridinyl-1,4-benzoquinone (DZQ). Labelled probes are then synthesized from the remaining single stranded cDNA species (unreacted mRNA species remaining from the driver are not converted into probe material due to specificity of Sequenase T7 DNA polymerase used to make the probe) and used to screen a cDNA library made from the tester cell population. A schematic diagram of the system is shown in figure 3.

It has been shown that the differentially expressed sequences can be enriched at least 300-fold with one round of subtraction (Hampson *et al.* 1992), and that the technique should allow isolation of cDNAs derived from transcripts that are present at less than 50 copies per cell. This equates to genes at the low end of intermediate abundance (see table 1). The main advantages of the CCLS approach are that it is rapid, technically simple and also produces fewer false positives than other differential expression analysis methods. However, like the physical separation protocols, a major drawback with CCLS is the large amount of starting material required (at least 10 μg RNA). Consequently, the technique has recently been refined so that a renewable source of RNA can be generated. The degenerate random oligonucleotide primed (DROP) adaptation (Hampson *et al.* 1996, Hampson and Hampson 1997) uses random hexanucleotide sequences to prime solid phase-synthesized cDNA. Since each primer includes a T7 polymerase promotor sequence

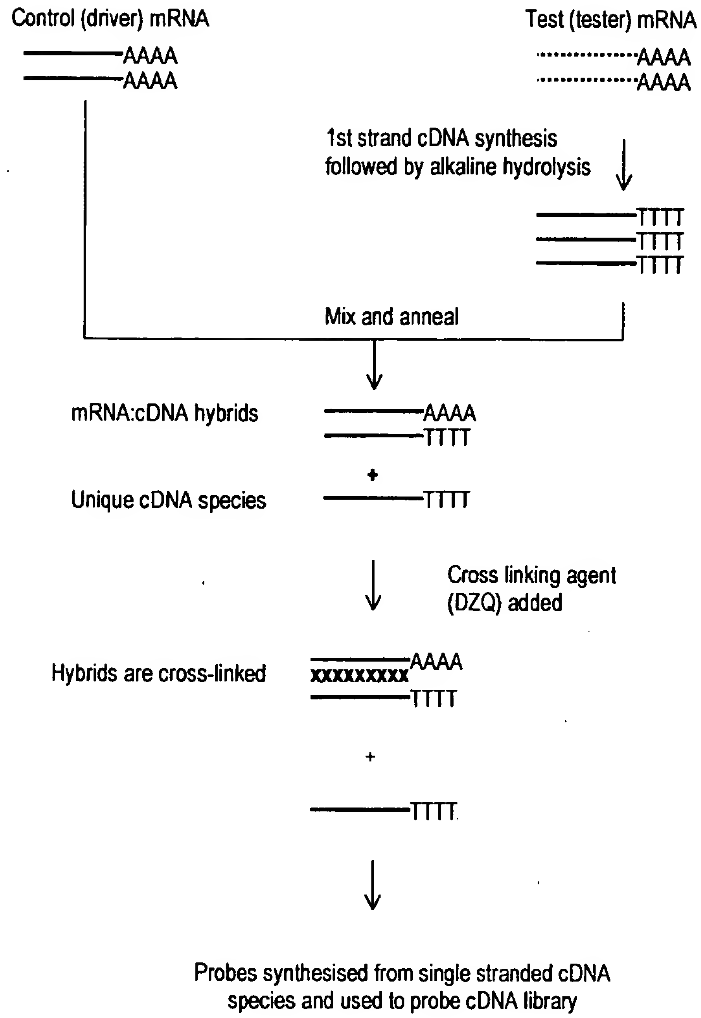


Figure 3. Chemical cross-linking subtraction. Excess driver mRNA is mixed with 1st strand tester cDNA. The common sequences form mRNA:cDNA hybrids which are cross linked with 2,5 diaziridinyl-1,4-benzoquinone (DZQ) and the remaining cDNA sequences are differentially expressed in the tester population. Probes are made from these sequences using Sequenase 2.0 DNA polymerase, which lacks reverse transcriptase activity and, therefore, does not react with the remaining mRNA molecules from the driver. The labelled probes are then used to screen a cDNA library for clones of differentially expressed sequences. Adapted from Walter *et al.* (1996), with permission.

Table 1. The abundance of mRNA species and classes in a typical mammalian cell.

mRNA class	Copies of each species/cell	No. of mRNA species in class	Mean % of each species in class	Mean mass (ng) of each species/μg total RNA
Abundant	12000	4	3.3	1.65
Intermediate	300	500	0.08	0.04
Rare	15	11000	0.004	0.002

Modified from Bertioli *et al.* (1995).

at the 5′ end, the final pool of random cDNA fragments is a PCR-renewable cDNA population which is representative of the expressed gene pool and can be used to synthesize sense RNA for use as driver material. Furthermore, if the final pool of random cDNA fragments is reamplified using biotinylated T7 primer and random hexamer, the product can be captured with streptavidin beads and the antisense strand eluted for use as tester. Since both target and driver can be generated from the same DROP product, subtraction can be performed in both directions (i.e. for up- and down-regulated species) between two different DROP products.

Representational Difference Analysis (RDA)

RDA of cDNA (Hubank and Schatz 1994) is an extension of the technique originally applied to genomic DNA as a means of identifying differences between two complex genomes (Lisitsyn *et al.* 1993). It is a process of subtraction and amplification involving subtractive hybridization of the tester in the presence of excess driver. Sequences in the tester that have homologues in the driver are rendered unamplifiable, whereas those genes expressed only in the tester retain the ability to be amplified by PCR. The procedure is shown schematically in figure 4.

In essence, the driver and tester mRNA populations are first converted to cDNA and amplified by PCR following the ligation of an adaptor. The adaptors are then removed from both populations and a new (different) adaptor ligated to the amplified tester population only. Driver and tester populations are next melted and hybridized together in a ratio of 100:1. Following hybridization, only tester:tester homohybrids have 5′ adaptors at each end of the DNA duplex and can, thus, be filled in at both 3′ ends. Hence, only these molecules are amplified exponentially during the subsequent PCR step. Although tester:driver heterohybrids are present, they only amplify in a linear fashion, since the strand derived from the driver has no adaptor to which the primer can bind. Driver:driver heterohybrids have no adaptors and, therefore, are not amplified. Single stranded molecules are digested with mung bean nuclease before a further PCR-enrichment of the tester:tester homohybrids. The adaptors on the amplified tester population are then replaced and the whole process repeated a further two or three times using an increasing excess of driver (Hubank and Shatz used a tester:driver ratio of 1:400, 1:80000 and 1:800000 for the second, third and fourth hybridizations, respectively). Different adaptors are ligated to the tester between successive rounds of hybridization and amplification to prevent the accumulation of PCR products that might interfere with subsequent amplifications. The final display is a series of differentially expressed gene products easily observable on an ethidium bromide gel.

The main advantages of RDA are that it offers a reproducible and sensitive approach to the analysis of differentially expressed genes. Hubank and Schatz (1994) reported that they were able to isolate genes that were differentially expressed in substantially less than 1% of the cells from which the tester is derived. Perhaps the main drawback is that multiple rounds of ligation, hybridization, amplifiation and digestion are required. The procedure is, therefore, lengthier than many other differential display approaches and provides more opportunity for operator-induced error to occur. Although the generation of false positives has been noted, this has been solved to some degree by O'Neill and Sinclair (1997) through the use of HPLC-purified adaptors. These are free of the truncated adaptors which appear to be a major source of the false positive bands. A very similar technique to RDA, termed linker capture subtraction (LCS) was described by Yang and Sytowski (1996).

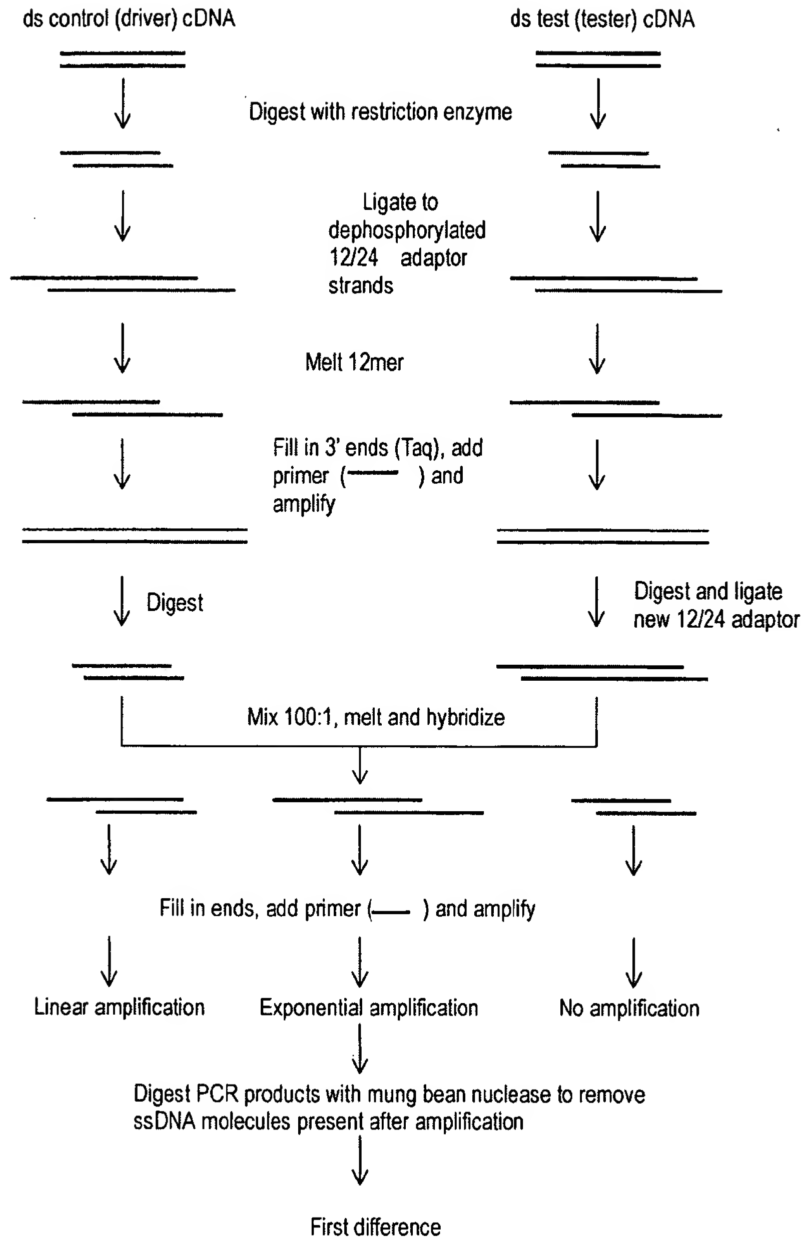


Figure 4. The representational difference analysis (RDA) technique. Driver and tester cDNA are digested with a 4-cutter restriction enzyme such as *Dpn*II. The 1st set of 12/24 adaptor strands (oligonucleotides) are ligated to each other and the digested cDNA products. The 12mer is subsequently melted away and the 3'ends filled in using Taq DNA polymerase. Each cDNA population is then amplified using PCR, following which the 1st set of adaptors is removed with *Dpn*II. A second set of 12/24 adaptor strands is then added to the amplified tester cDNA population, after which the tester is hybridized against a large excess of driver. The 12mer adaptors are melted and the 3′ ends filled in as before. PCR is carried out with primers identical to the new 24mer adaptor. Thus, the only hybridization products which are exponentially amplified are those which are tester:tester combinations. Following PCR, ssDNA products are removed with mung bean nuclease, leaving the 'first difference product'. This is digested and a third set of 12/24 adaptors added before repeating the subtraction process from the hybridization stage. The process is repeated to the 3rd or 4th difference product, as described by Lisitsyn *et al.* (1993) and Hubank and Schatz (1994).

Suppression PCR Subtractive Hybridization (SSH)

The most recent adaptation of the SH approach to differential expression analysis was first described by Diatchenko *et al.* (1996) and Gurskaya *et al.* (1996). They reported that a 1000–5000 fold enrichment of rare cDNAs (equivalent to isolating mRNAs present at only a few copies per cell) can be obtained without the need for multiple hybridizations/subtractions. Instead of physical or chemical removal of the common sequences, a PCR-based suppression system is used (see figure 5).

In SSH, excess driver cDNA is added to two portions of the tester cDNA which have been ligated with different adaptors. A first round of hybridization serves to enrich differentially expressed genes and equalize rare and abundant messages. Equalization occurs since reannealing is more rapid for abundant molecules than for rarer molecules due to the second order kinetics of hybridization (James and Higgins 1985). The two primary hybridization mixes are then mixed together in the presence of excess driver and allowed to hybridize further. This step permits the annealing of single stranded complementary sequences which did not hybridize in the primary hybridization, and in doing so generates templates for PCR amplification. Although there are several possible combinations of the single stranded molecules present in the secondary hybridization mix, only one particular combination (differentially expressed in the tester cDNA composed of complimentary strands having different adaptors) can amplify exponentially.

Having obtained the final differential display, two options are available if cloning of cDNAs is desired. One is to transform the whole of the final PCR reaction into competent cells. Transformed colonies can then be isolated and their inserts characterized by sequencing, restriction analysis or PCR. Alternatively, the final PCR products can be resolved on a gel and the individual bands excised, reamplified and cloned. The first approach is technically simpler and less time consuming. However, ligation/transformation reactions are known to be biased towards the cloning of smaller molecules, and so the final population of clones will probably not contain a representative selection of the larger products. In addition, although equalization theoretically occurs, observations in this laboratory suggest that this is by no means perfectly accomplished. Consequently, some gene species are present in a higher number than others and this will be represented in the final population of clones. Thus, in order to obtain a substantial proportion of those gene species that actually demonstrate differential expression in the tester population, the number of clones that will have to be screened after this step may be substantial. The second approach is initially more time consuming and technically demanding. However, it would appear to offer better prospects for cloning larger and low abundance gel products. In addition, one can incorporate a screening step that differentiates different products of different sequences but of the same size (HA-staining, see later). In this way, a good idea of the final number of clones to be isolated and identified can be achieved.

An alternative (or even complementary) approach is to use the final differential display reaction to screen a cDNA library to isolate full length clones for further characterization, or a DNA array (see later) to quickly identify known genes. SSH has been used in this laboratory to begin characterization of the short-term gene expression profiles of enzyme-inducers such as phenobarbital (Rockett *et al.* 1997) and Wy-14,643 (Rockett *et al.* unpublished observations). The isolation of differentially expressed genes in this manner enables the construction of a fingerprint

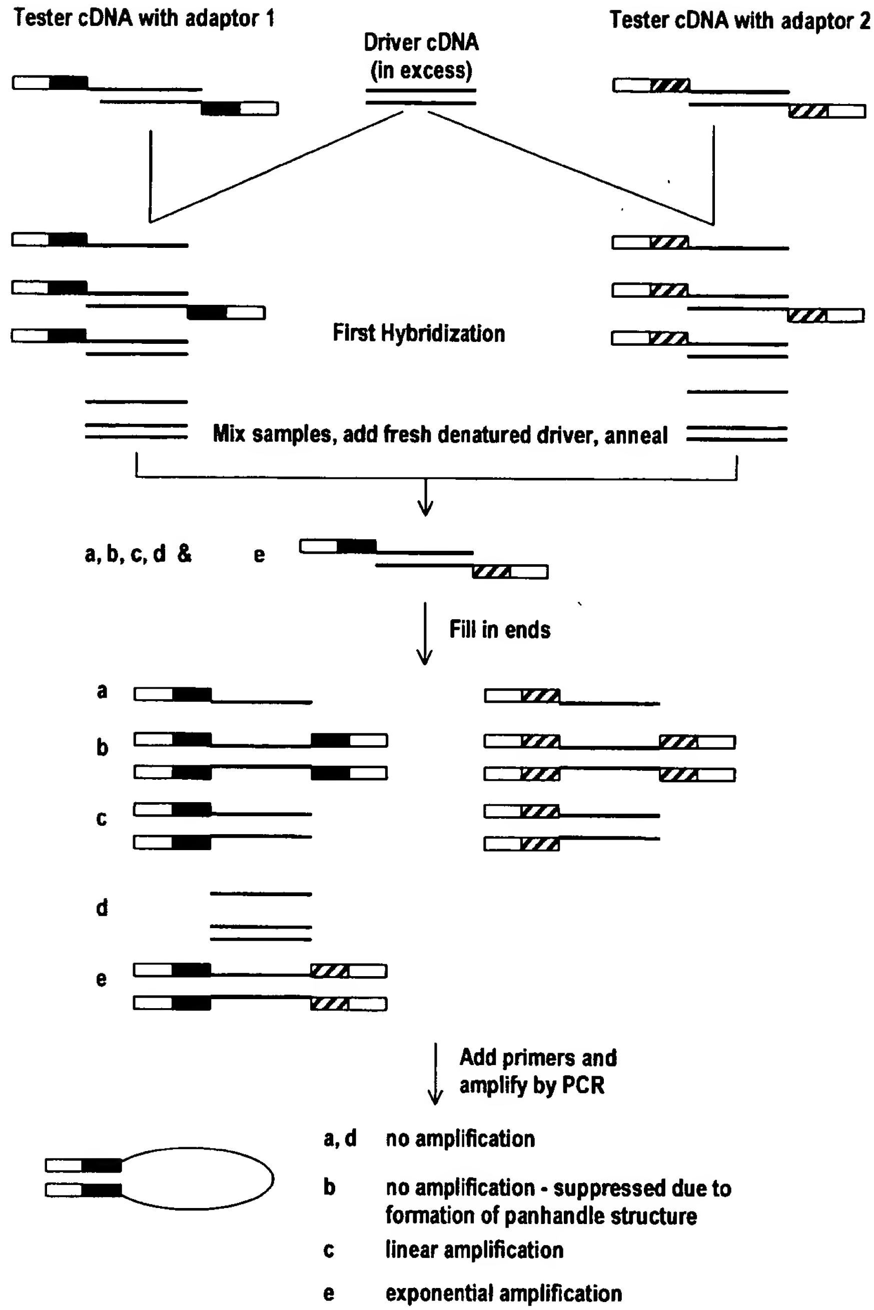


Figure 5. PCR-select cDNA subtraction. In the primary hybridization, an excess of driver cDNA is added to each tester cDNA population. The samples are heat denatured and allowed to hybridize for between 3 and 8 h. This serves two purposes: (1) to equalize rare and abundant molecules; and (2) to enrich for differentially expressed sequences—cDNAs that are not differentially expressed form type c molecules with the driver. In the secondary hybridization, the two primary hybridizations are mixed together without denaturing. Fresh denatured driver can also be added at this point to allow further enrichment of differentially expressed sequences. Type e molecules are formed in this secondary hybridization which are subsequently amplified using two rounds of PCR. The final products can be visualized on an agarose gel, labelled directly or cloned into a vector for downstream manipulation. As described by Diatchenko *et al.* (1996) and Gurskaya *et al.* (1996), with permission.

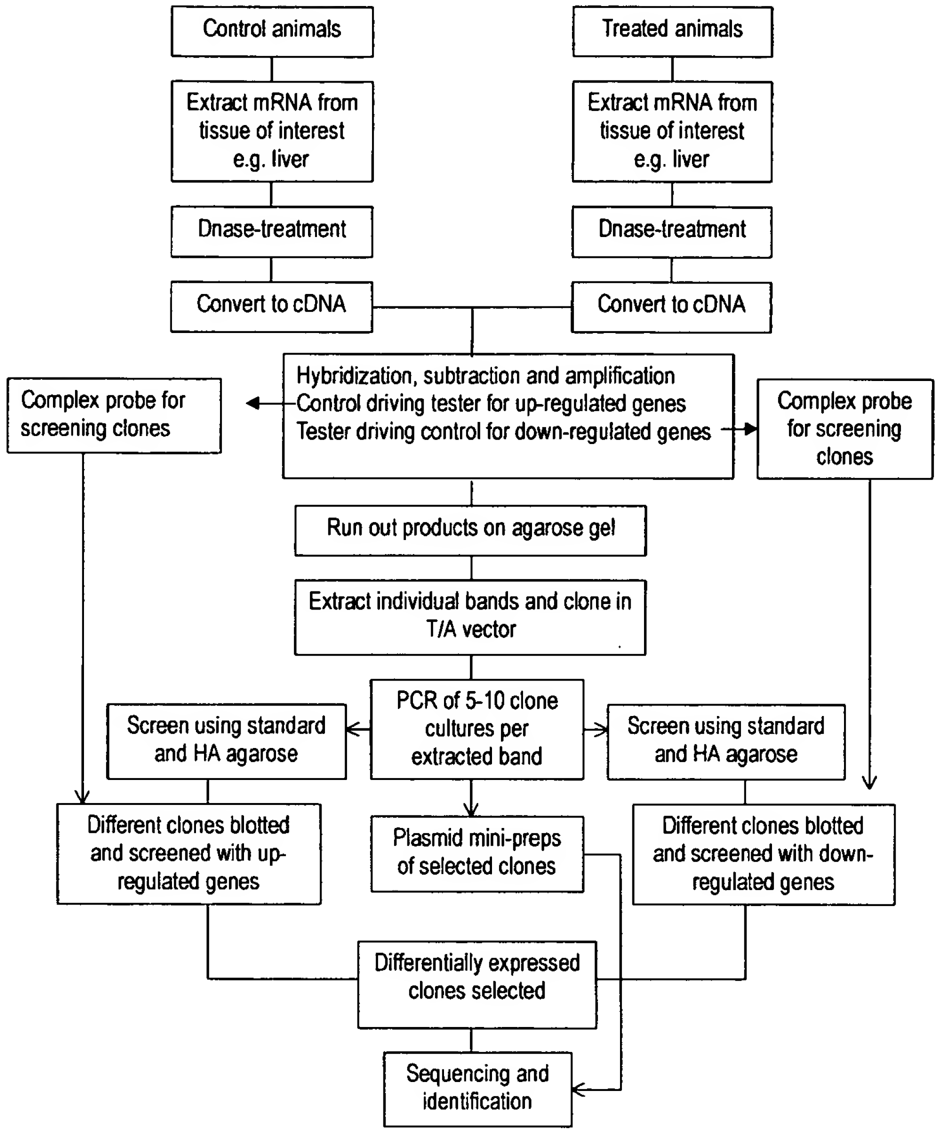


Figure 6. Flow diagram showing method used in this laboratory to isolate and identify clones of genes which are differentially expressed in rat liver following short term exposure to the enzyme inducers, phenobarbital and Wy-14,643.

of expressed genes which are unique to each compound and time/dose point. Such information could be useful in short-term characterization of the toxic potential of new compounds by comparing the gene-expression profiles they elicit with those produced by known inducers. Figure 6 shows a flow diagram of the method used to isolate, verify and clone differentially expressed genes, and figure 7 shows expression profiles obtained from a typical SSH experiment. Subsequent sub-cloning of the individual bands, sequencing and gene data base interrogation reveals many genes which are either up- or down-regulated by phenobarbital in the rat (tables 2 and 3).

One of the advantages in using the SSH approach is that no prior knowledge is required of which specific genes are up/down-regulated subsequent to xenobiotic

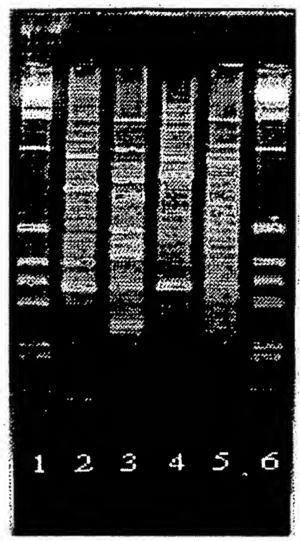


Figure 7. SSH display patterns obtained from rat liver following 3-day treatment with WY-14,643 or phenobarbital. mRNA extracted from control and treated livers was used to generate the differential displays using the PCR-Select cDNA subtraction kit (Clontech). Lane: 1—1kb ladder; 2—genes upregulated following Wy,14–643 treatment; 3—genes downregulated following Wy,14–643 treatment; 4—genes upregulated following phenobarbital treatment; 5—genes downregulated following phenobarbital treatment; 6—1kb ladder. Reproduced from Rockett *et al.* (1997), with permission.

exposure, and an almost complete complement of genes are obtained. For example, the peroxisome proliferator and non-genotoxic hepatocarcinogen Wy,14,643, up-regulates at least 28 genes and down-regulates at least 15 in the rat (a sensitive species) and produces 48 up- and 37 down-regulated genes in the guinea pig, a resistant species (Rockett, Swales, Esda and Gibson, unpublished observations). One of these genes, CD81, was up-regulated in the rat and down-regulated in the guinea pig following Wy-14,643 treatment. CD81 (alternatively named TAPA-1) is a widely expressed cell surface protein which is involved in a large number of cellular processes including adhesion, activation, proliferation and differentiation (Levy *et al.* 1998). Since all of these functions are altered to some extent in the phenomena of hepatomegaly and non-genotoxic hepatocarcinogenesis, it is intriguing, and probably mechanistically-relevant, that CD81 expression is differentially regulated in a resistant and susceptible species. However, the down-side of this approach is that the majority of genes can be sequenced and matched to database sequences, but the latter are predominantly expressed sequence tags or genes of completely unknown function, thus partially obscuring a realistic overall assessment of the critical genes of genuine biological interest. Notwithstanding the lack of complete funtional identification of altered gene expression, such gene profiling studies essentially provides a 'molecular fingerprint' in response to xenobiotic challenge, thereby serving as a mechanistically-relevant platform for further detailed investigations.

Differential Display (DD)

Originally described as 'RNA fingerprinting by arbitrarily primed PCR' (Liang and Pardee 1992) this method is now more commonly referred to as 'differential

Table 2. Genes up-regulated in rat liver following 3-day exposure to phenobarbital.

Band number (approximate size in bp)	Highest sequence similarity	FASTA-EMBL gene identification
5 (1300)	93.5%	CYP2B1
7 (1000)	95.1%	Preproalbumin
		Serum albumin mRNA
8 (950)	98.3%	NCI-CGAP-Pr1 *H. sapiens* (EST)
10 (850)	95.7%	CYP2B1
11 (800)	Clone 1 94.9%	CYP2B1
	Clone 2 75.3%	CYP2B2
12 (750)	93.8%	TRPM-2 mRNA
		Sulfated glycoprotein
15 (600)	92.9%	Preproalbumin
		Serum albumin mRNA
16 (55)	Clone 1 95.2%	CYP2B1
	Clone 2 93.6%	Haptoglobulin mRNA partial alpha
21 (350)	99.3%	18S, 5.8S & 28S rRNa

Bands 1–4, 6, 9, 13, 14, and 17–20 are shown to be false positives by dot blot anaylsis and, therefore, are not sequenced. Derived from Rockett *et al.* (1997). It should be noted that the above genes do not represent the complete spectrum of genes which are up-regulated in rat liver by phenobarbital, but simply represents the genes sequenced and identified to date.

Table 3. Genes down-regulated in rat liver following 3-day exposure to phenobarbital.

Band number (approximate size in bp)	Highest sequence similarity	FASTA-EMBL gene identification
1 (1500)	95.3%	3-oxoacyl-CoA thiolase
2 (1200)	92.3%	Hemopoxin mRNA
3 (1000)	91.7%	Alpha-2u-globulin mRNA
7 (700)	Clone 1 77.2%	*M.musculus* Cl inhibitor
	Clone 2 94.5%	Electron transfer flavoprotein
	Clone 3 91.0%	*M. musculus* Topoisomerase 1 (Topo 1)
8 (650)	Clone 1 86.9%	Soares 2NbMT *M. musculus* (EST)
	Clone 2 96.2%	Alpha-2u-globulin (s-type) mRNA
9 (600)	Clone 1 86.9%	Soares mouse NML *M. musculus* (EST)
	Clone 2 82.0%	Soares p3NMF 19.5 *M. musculus* (EST)
10 (550)	73.8%	Soares mouse NML *M. musculus* (EST)
11 (525)	95.7%	NCI-CGAP-Pr1 *H. sapiens* (EST)
12 (375)	100.0%	Ribosomal protein
13 (23)	Clone 1 97.2%	Soares mouse embryo NbME135 (EST)
	Clone 2 100.0%	Fibrinogen B-beta-chain
	Clone 3 100.0%	Apolipoprotein E gene
14 (170)	96.0%	Soares p3NMF19.5 *M. musculus* (EST)
15 (140)	97.3%	Stratagene mouse testis (EST)
Others: (300)	96.7%	*R. norvegicus* RASP 1 mRNA
(275)	93.1%	Soares mouse mammary gland (EST)

EST = Expressed sequence tag. Bands 4–6 were shown to be false positives by dot blot analysis and, therefore, were not sequenced. Derived from Rockett *et al.* (1997). It should be noted that the above genes do not represent the complete spectrum of genes which are down-regulated in rat liver by phenobarbital, but simiply represents the genes sequenced and identified to date.

display' (DD). In this method, all the mRNA species in the control and treated cell populations are amplified in separate reactions using reverse transcriptase-PCR (RT-PCR). The products are then run side-by-side on sequencing gels. Those bands which are present in one display only, or which are much more intense in one

display compared to the other, are differentially expressed and may be recovered for further characterization. One advantage of this system is the speed with which it can be carried out—2 days to obtain a display and as little as a week to make and identify clones.

Two commonly used variations are based on different methods of priming the reverse transcription step (figure 8). One is to use an oligo dT with a 2-base 'anchor' at the 3'-end, e.g. 5′ (dT_{11})CA 3′ (Liang and Pardee 1992). Alternatively, an arbitrary primer may be used for 1st strand cDNA synthesis (Welsh *et al.* 1992). This variant of RNA fingerprinting has also been called 'RAP' (RNA Arbitrarily Primed)-PCR. One advantage of this second approach is that PCR products may be derived from anywhere in the RNA, including open reading frames. In addition, it can be used for mRNAs that are not polyadenylated, such as many bacterial mRNAs (Wong and McClelland 1994). In both cases, following reverse transcription and denaturation, second strand cDNA synthesis is carried out with an arbitrary primer (*arbitrary* primers have a single base at each position, as compared to *random* primers, which contain a mixture of all four bases at each position). The resulting PCR, thus, produces a series of products which, depending on the system (primer length and composition, polymerase and gel system), usually includes 50–100 products per primer set (Band and Sager 1989). When a combination of different dT-anchors and arbitrary primers are used, almost all mRNA species from a cell can be amplified. When the cDNA products from two different populations are analysed side by side on a polyacrylamide gel, differences in expression can be identified and the appropriate bands recovered for cloning and further analysis.

Although DD is perhaps the most popular approach used today for identifying differentially expressed genes, it does suffer from several perceived disadvantages:

(1) It may have a strong bias towards high copy number mRNAs (Bertioli *et al.* 1995), although this has been disputed (Wan *et al.* 1996) and the isolation of very low abundance genes may be achieved in certain circumstances (Guimeraes *et al.* 1995a).
(2) The cDNAs obtained often only represent the extreme 3′ end of the mRNA (often the 3′-untranslated region), although this may not always be the case (Guimeraes *et al.* 1995a). Since the 3′ end is often not included in Genbank and shows variation between organisms, cDNAs identified by DD cannot always be matched with their genes, even if they have been identified.
(3) The pattern of differential expression seen on the display often cannot be reproduced on Northern blots, with false positives arising in up to 70% of cases (Sun *et al.* 1994). Some adaptations have been shown to reduce false positives, including the use of two reverse transcriptases (Sung and Denman 1997), comparison of uninduced and induced cells over a time course (Burn *et al.* 1994) and comparison of DDPCR-products from two uninduced and two induced lines (Sompayrac *et al.* 1995). The latter authors also reported that the use of cytoplasmic RNA rather then total RNA reduces false positives arising from nuclear RNA that is not transported to the cytoplasm.

Further details of the background, strengths and weaknesses of the DD technique can be obtained from a review by McClelland *et al.* (1996) and from articles by Liang *et al.* (1995) and Wan *et al.* (1996).

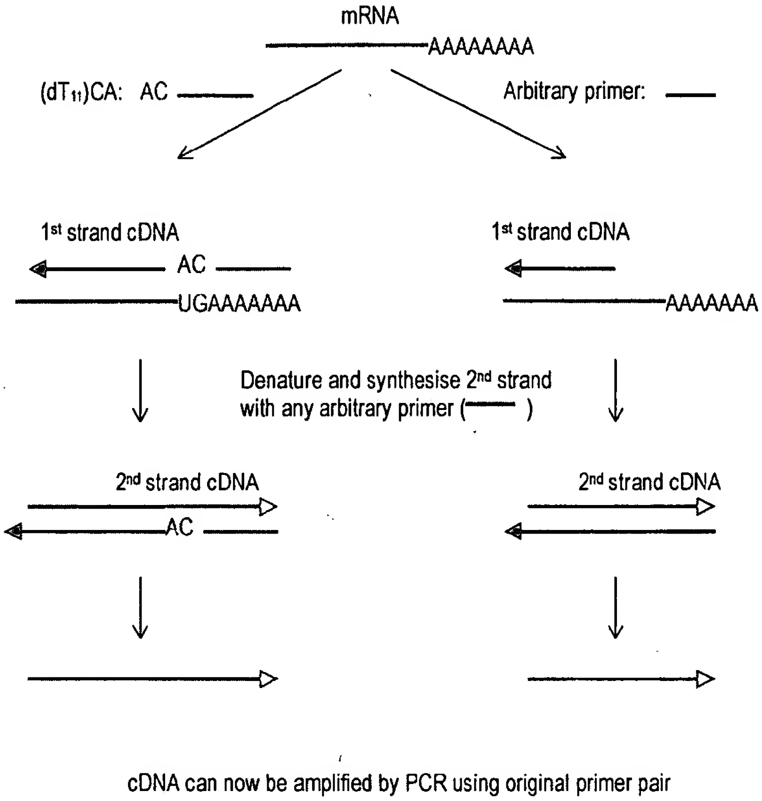


Figure 8. Two approaches to differential display (DD) analysis. 1st strand synthesis can be carried out either with a polydT$_{11}$NN primer (where N = G, C or A) or with an arbitrary primer. The use of different combinations of G, C and A to anchor the first strand polydT primer enables the priming of the majority of polyadenylated mRNAs. Arbitrary primers may hybridize at none, one or more places along the length of the mRNA, allowing 1st strand cDNA synthesis to occur at none, one or more points in the same gene. In both cases, 2nd strand synthesis is carried out with an arbitrary primer. Since these arbitrary primers for the 2nd strand may also hybridize to the 1st strand cDNA in a number of different places, several different 2nd strand products may be obtained from one binding point of the 1st strand primer. Following 2nd strand synthesis, the original set of primers is used to amplify the second strand products, with the result that numerous gene sequences are amplified.

Restriction endonuclease-facilitated analysis of gene expression

Serial Analysis of Gene Expression (SAGE)

A more recent development in the field of differential display is SAGE analysis (Velculescu *et al.* 1995). This method uses a different approach to those discussed so far and is based on two principles. Firstly, in more than 95% of cases, short nucleotide sequences ('tags') of only nine or 10 base pairs provide sufficient information to identify their gene of origin. Secondly, concatonation (linking together in a series) of these tags allows sequencing of multiple cDNAs within a single clone. Figure 9 shows a schematic representation of the SAGE process. In this procedure, double stranded cDNA from the test cells is synthesized with a biotinylated polydT primer. Following digestion with a commonly cutting (4bp recognition sequence) restriction enzyme ('anchoring enzyme'), the 3′ ends of the cDNA population are captured with streptavidin beads. The captured population is

split into two and different adaptors ligated to the 5′ ends of each group. Incorporated into the adaptors is a recognition sequence for a type IIS restriction enzyme—one which cuts DNA at a defined distance (< 20 bp) from its recognition sequence. Hence, following digestion of each captured cDNA population with the IIS enzyme, the adaptors plus a short piece of the captured cDNA are released. The two populations are then ligated and the products amplified. The amplified products are cleaved with the original anchoring enzyme, religated (concatomers are formed in the process) and cloned. The advantage of this system is that hundreds of gene tags can be identified by sequencing only a few clones. Furthermore, the number of times a given transcript is identified is a quantitative measurement of that gene's abundance in the original population, a feature which facilitates identification of differentially expressed genes in different cell populations.

Some disadvantages of SAGE analysis include the technical difficulty of the method, a large amount of accurate sequencing is required, biased towards abundant mRNAs, has not been validated in the pharmaco/toxicogenomic setting and has only been used to examine well known tissue differences to date.

Gene Expression Fingerprinting (GEF)

A different capture/restriction digest approach for isolating differentially expressed genes has been described by Ivanova and Belyavsky (1995). In this method, RNA is converted to cDNA using biotinylated oligo(dT) primers. The cDNA population is then digested with a specific endonuclease and captured with magnetic streptavidin microbeads to facilitate removal of the unwanted 5′ digestion products. The use of restricted 3′-ends alone serves to reduce the complexity of the cDNA fragment pool and helps to ensure that each RNA species is represented by not more than one restriction product. An adaptor is ligated to facilitate subsequent amplification of the captured population. PCR is carried out with one adaptor-specific and one biotinylated polydT primer. The reamplified population is recaptured and the non-biotinylated strands removed by alkaline dissociation. The non-biotinylated strand is then resynthesized using a different adaptor-specific primer in the presence of a radiolabelled dNTP. The labelled immobilized 3′ cDNA ends are next sequentially treated with a series of different restriction endonucleases and the products from each digestion analysed by PAGE. The result is a fingerprint composed of a number of ladders (equal to the number of sequential digests used). By comparing test versus control fingerprints, it is possible to identify differentially expressed products which can then be isolated from the gel and cloned. The advantages of this procedure are that it is very robust and reproducible, and the authors estimate that 80–93% of cDNA molecules are involved in the final fingerprint. The disadvantage is that polyacrylamide gels can rarely resolve more than 300–400 bands, which compares poorly to the 1000 or more which are estimated to be produced in an average experiment. The use of 2-D gels such as those described by Uitterlinden *et al.* (1989) and Hatada *et al.* (1991) may help to overcome this problem.

A similar method for displaying restriction endonuclease fragments was later described by Prashar and Weissman (1996). However, instead of sequential digestion of the immobolized 3'-terminal cDNA fragments, these authors simply compared the profiles of the control and treated populations without further manipulation.

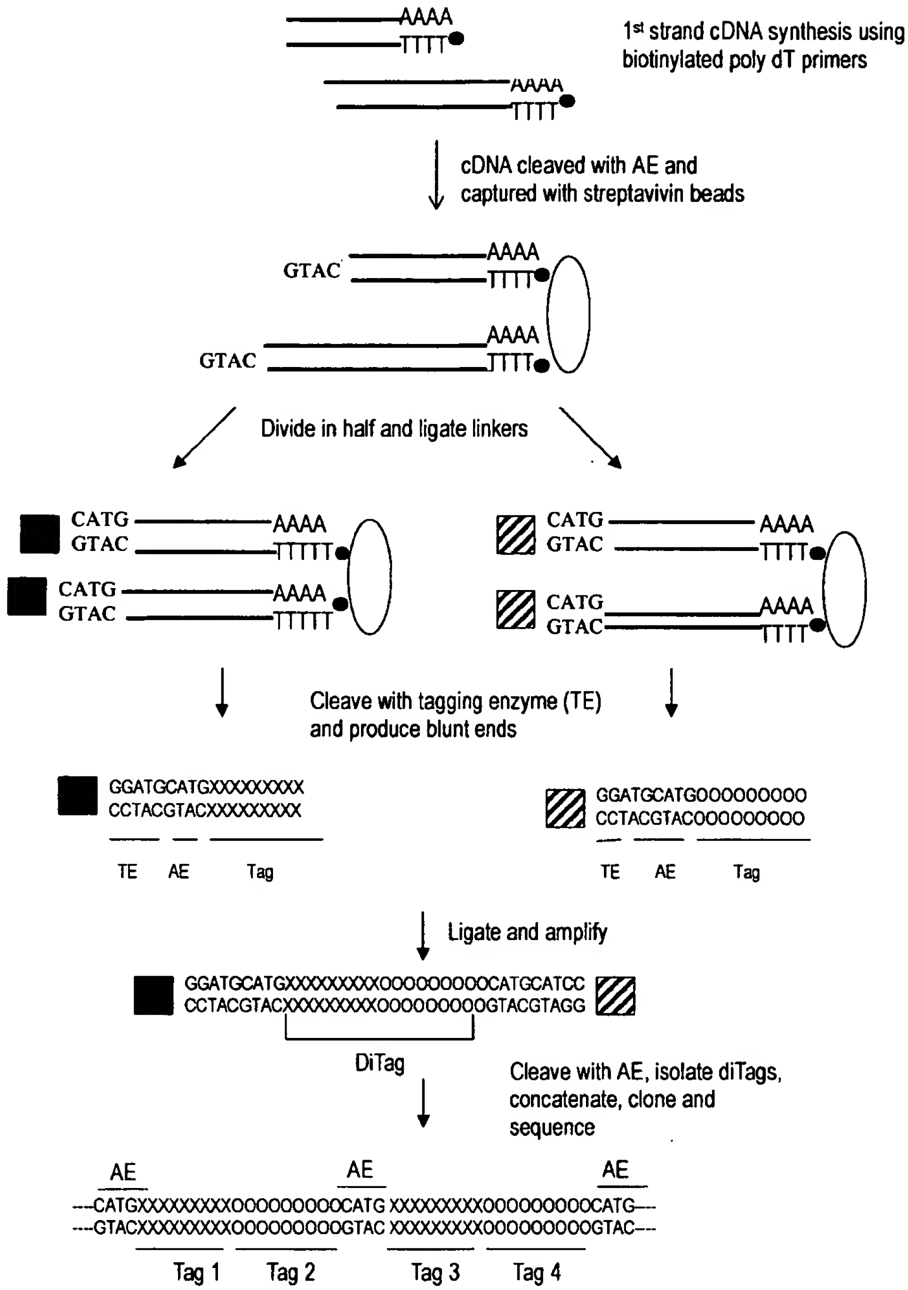


Figure 9. Serial analysis of gene expression (SAGE) analysis. cDNA is cleaved with an anchoring enzyme (AE) and the 3'ends captured using streptavidin beads. The cDNA pool is divided in half and each portion ligated to a different linker, each containing a type IIS restriction site (tagging enzyme, TE). Restriction with the type IIS enzyme releases the linker plus a short length of cDNA (XXXXX and OOOOO indicate nucleotides of different tags). The two pools of tags are then ligated and amplified using linker-specific primers. Following PCR, the products are cleaved with the AE and the ditags isolated from the linkers using PAGE. The ditags are then ligated (during which process, concatenization occurs) and cloned into a vector of choice for sequencing. After Velculescu *et al.* (1995), with permission.

DNA arrays

'Open' differential display systems are cumbersome in that it takes a great deal of time to extract and identify candidate genes and then confirm that they are indeed up- or down-regulated in the treated compared to the control tissue. Normally, the latter process is carried out using Northern blotting or RT-PCR. Even so, each of the aforementioned steps produce a bottleneck to the ultimate goal of rapid analysis of gene expression. These problems will likely be addressed by the development of so-called DNA arrays (e.g. Gress *et al.* 1992, Zhao *et al.* 1995, Schena *et al.* 1996), the introduction of which has signalled the next era in differential gene expression analysis. DNA arrays consist of a gridded membrane or glass 'chips' containing hundreds or thousands of DNA spots, each consisting of multiple copies of part of a known gene. The genes are often selected based on previously proven involvement in oncogenesis, cell cycling, DNA repair, development and other cellular processes. They are usually chosen to be as specific as possible for each gene and animal species. Human and mouse arrays are already commercially available and a few companies will construct a personalized array to order, for example Clontech Laboratories and Research Genetics Inc. The technique is rapid in that hundreds or even thousands of genes can be spotted on a single array, and that mRNA/cDNA from the test populations can be labelled and used directly as probe. When analysed with appropriate hardware and software, arrays offer a rapid and quantitative means to assess differences in gene expression between two cell populations. Of course, there can only be identification and quantitation of those genes which are in the array (hence the term 'closed' system). Therefore, one approach to elucidating the molecular mechanisms involved in a particular disease/development system may be to combine an open and closed system—a DNA array to directly identify and quantitate the expression of known genes in mRNA populations, and an open system such as SSH to isolate unknown genes which are differentially expressed.

One of the main advantages of DNA arrays is the huge number of gene fragments which can be put on a membrane—some companies have reported gridding up to 60000 spots on a single glass 'chip' (microscope slide). These high density chip-based micro-arrays will probably become available as mass-produced off-the-shelf items in the near future. This should facilitate the more rapid determination of differential expression in time and dose-response experiments. Aside from their high cost and the technical complexities involved in producing and probing DNA arrays, the main problem which remains, especially with the newer micro-array (gene-chip) technologies, is that results are often not wholly reproducible between arrays. However, this problem is being addressed and should be resolved within the next few years.

EST databases as a means to identify differentially expressed genes

Expressed sequence tags (ESTs) are partial sequences of clones obtained from cDNA libraries. Even though most ESTs have no formal identity (putative identification is the best to be hoped for), they have proven to be a rapid and efficient means of discovering new genes and can be used to generate profiles of gene-expression in specific cells. Since they were first described by Adams *et al.* (1991), there has been a huge explosion in EST production and it is estimated that there are now well over a million such sequences in the public domain, representing over half

of all human genes (Hillier *et al.* 1996). This large number of freely available sequences (both sequence information and clones are normally available royalty-free from the originators) has enabled the development of a new approach towards differential gene expression analysis as described by Vasmatzis *et al.* (1998). The approach is simple in theory: EST databases are first searched for genes that have a number of related EST sequences from the target tissue of choice, but none or few from non-target tissue libraries. Programmes to assist in the assembly of such sets of overlapping data may be developed in-house or obtained privately or from the internet. For example, the Institute for Genomic Research (TIGR, found at http://www.tigr.org) provides many software tools free of charge to the scientific community. Included amongst these is the TIGR assembler (Sutton *et al.* 1995), a tool for the assembly of large sets of overlapping data such as ESTs, bacterial artificial chromosomes (BAC)s, or small genomes. Candidate EST clones representing different genes are then analysed using RNA blot methods for size and tissue specificity and, if required, used as probes to isolate and identify the full length cDNA clone for further characterization. In practice however, the method is rather more involved, requiring bioinformatic and computer analysis coupled with confirmatory molecular studies. Vasmatzis *et al.* (1998) have described several problems in this fledgling approach, such as separating highly homologous sequences derived from different genes and an overemphasis of specificity for some EST sequences. However, since these problems will largely be addressed by the development of more suitable computer algorithms and an increased completeness of the EST database, it is likely that this approach to identifying differentially expressed genes may enjoy more patronage in the future.

Problems and potential of differential expression techniques

The holistic or single cell approach?

When working with *in vivo* models of differential expression, one of the first issues to consider must be the presence of multiple cell types in any given specimen. For example, a liver sample is likely to contain not only hepatocytes, but also (potentially) Ito cells, bile ductule cells, endothelial cells, various immune cells (e.g. lymphocytes, macrophages and Kupffer cells) and fibroblasts. Other tissues will each have their own distinctive cell populations. Also, in the case of neoplastic tissue, there are almost always normal, hyperplastic and/or dysplastic cells present in a sample. One must, therefore, be aware that genes obtained from a differential display experiment performed on an animal tissue model may not necessarily arise exclusively from the intended 'target' cells, e.g. hepatocytes/neoplastic cells. If appropriate, further analyses using immunohistochemistry, *in situ* hybridization or *in situ* RT-PCR should be used to confirm which cell types are expressing the gene(s) of interest. This problem is probably most acute for those studying the differential expression of genes in the development of different cell types, where there is a need to examine homologous cell populations. The problem is now being addressed at the National Cancer Institute (Bethesda, MD, USA) where new microdisection techniques have been employed to assist in their gene analysis programme, the Cancer Genome Anatomy Project (CGAP) (For more information see web site: http://www.ncbi.nlm.nih.gov/ncicgap/intro.html). There are also separation techniques available that utilise cell-specific antigens as a means to isolate target cells,

e.g. fluorescence activated cell sorting (FACS) (Dunbar *et al.* 1998, Kas-Deelen *et al.* 1998) and magnetic bead technology (Richard *et al.* 1998, Rogler *et al.* 1998).

However, those taking a holistic approach may consider this issue unimportant. There is an equally appropriate view that all those genes showing altered expression within a compromized tissue should be taken into consideration. After all, since all tissues are complex mixes of different, interacting cell types which intimately regulate each other's growth and development, it is clear that each cell type could in some way contribute (positively or negatively) towards the molecular mechanisms which lie behind responses to external stimuli or neoplastic growth. It is perhaps then more informative to carry out differential display experiments using *in vivo* as opposed to *in vitro* models, where uniform populations of identical cells probably represent a partial, skewed or even inaccurate picture of the molecular changes that occur.

The incidence and possible implications of inter-individual biological variation should be considered in any approach where whole animal models are being used. It is clear that individuals (humans and animals) respond in different ways to identical stimuli. One of the best characterized examples is the debrisoquine oxidation polymorphism, which is mediated by cytochrome CYP2D6 and determines the pharmacokinetics of many commonly prescribed drugs (Lennard 1993, Meyer and Zanger 1997). The reasons for such differences are varied and complex, but allelic variations, regulatory region polymorphisms and even physical and mental health can all contribute to observed differences in individual responses. Careful thought should, therefore, be given to the specific objectives of the study and to the possible value of pooling starting material (tissue/mRNA). The effect of this can be beneficial through the ironing out of exaggerated responses and unimportant minor fluctuations of (mechanistically) irrelevant genes in individual animals, thus providing a clearer overall picture of the general molecular mechanisms of the response. However, at the same time such minor variations may be of utmost importance in deciding the ability of individual animals to succumb to or resist the effects of a given chemical/disease.

How efficient are differential expression techniques at recovering a high percentage of differentially expressed genes?

A number of groups have produced experimental data suggesting that mammalian cells produce between 8000–15 000 different mRNA species at any one time (Mechler and Rabbitts 1981, Hedrick *et al.* 1984, Bravo 1990), although figures as high as 20–30 000 have also been quoted (Axel *et al.* 1976). Hedrick *et al.* (1984) provided evidence suggesting that the majority of these belong to the rare abundance class. A breakdown of this abundance distribution is shown in table 1.

When the results of differential display experiments have been compared with data obtained previously using other methods, it is apparent that not all differentially expressed mRNAs are represented in the final display. In particular, rare messages (which, importantly, often include regulatory proteins) are not easily recovered using differential display systems. This is a major shortcoming, as the majority of mRNA species exist at levels of less than 0.005% of the total population (table 1). Bertioli *et al.* (1995) examined the efficiency of DD templates (heterogeneous mRNA populations) for recovering rare messages and were unable to detect mRNA

species present at less than 1.2% of the total mRNA population—equivalent to an intermediate or abundant species. Interestingly, when simple model systems (single target only) were used instead of a heterogeneous mRNA population, the same primers could detect levels of target mRNA down to 10000× smaller. These results are probably best explained by competition for substrates from the many PCR products produced in a DD reaction.

The numbers of differentially expressed mRNAs reported in the literature using various model systems provides further evidence that many differentially expressed mRNAs are not recovered. For example, DeRisi *et al.* (1997) used DNA array technology to examine gene expression in yeast following exhaustion of sugar in the medium, and found that more than 1700 genes showed a change in expression of at least 2-fold. In light of such a finding, it would not be unreasonable to suggest that of the 8000–15 000 different mRNA species produced by any given mammalian cell, up to 1000 or more may show altered expression following chemical stimulation. Whilst this may be an extreme figure, it is known that at least 100 genes are activated/upregulated in Jurkat (T-) cells following IL-2 stimulation (Ullman *et al.* 1990). In addition, Wan *et al.* (1996) estimated that interferon-γ-stimulated HeLa cells differentially express up to 433 genes (assuming 24000 distinct mRNAs expressed by the cells). However, there have been few publications documenting anywhere near the recovery of these numbers. For example, in using DD to compare normal and regenerating mouse liver, Bauer *et al.* (1993) found only 70 of 38000 total bands to be different. Of these, 50% (35 genes) were shown to correspond to differentially expressed bands. Chen *et al.* (1996) reported 10 genes upregulated in female rat liver following ethinyl estradiol treatment. McKenzie and Drake (1997) identified 14 different gene products whose expression was altered by phorbol myristate acetate (PMA, a tumour promoter agent) stimulation of a human myelomonocytic cell line. Kilty and Vickers (1997) identified 10 different gene products whose expression was upregulated in the peripheral blood leukocytes of allergic disease sufferers. Linskens *et al.* (1995) found 23 genes differentially expressed between young and senescent fibroblasts. Techniques other than DD have also provided an apparent paucity of differentially expressed genes. Using SH for example, Cao *et al.* (1997) found 15 genes differentially expressed in colorectal cancer compared to normal mucosal epithelium. Fitzpatrick *et al.* (1995) isolated 17 genes upregulated in rat liver following treatment with the peroxisome proliferator, clofibrate; Philips *et al.* (1990) isolated 12 cDNA clones which were upregulated in highly metastatic mammary adenocarcinoma cell lines compared to poorly metastatic ones. Prashar and Weissman (1996) used 3′ restriction fragment analysis and identified approximately 40 genes showing altered expression within 4 h of activation of Jurkat T-cells. Groenink and Leegwater (1996) analysed 27 gene fragments isolated using SSH of delayed early response phase of liver regeneration and found only 12 to be upregulated.

In the laboratory, SSH was used to isolate up to 70 candidate genes which appear to show altered expression in guinea pig liver following short-term treatment with the peroxisome proliferator, WY-14,643 (Rockett, Swales, Esdaile and Gibson, unpublished observations). However, these findings have still to be confirmed by analysis of the extracted tissue mRNA for differential expression of these sequences.

Whilst the latest differential display technologies are purported to include design and experimental modifications to overcome this lack of efficiency (in both the total number of differentially expressed genes recovered and the percentage that are true

positives), it is still not clear if such adaptations are practically effective—proving efficiency by spiking with a known amount of limited numbers of artificial construct(s) is one thing, but isolating a high percentage of the rare messages already present in an mRNA population is another. Of course, some models will genuinely produce only a small number of differentially expressed genes. In addition, there are also technical problems that can reduce efficiency. For example, mRNAs may have an unusual primary structure that effectively prevents their amplification by PCR-based systems. In addition, it is known that under certain circumstances not all mRNAs have 3′polyA sites. For example, during *Xenopus* development, deadenylation is used as a means to stabilize RNAs (Voeltz and Steitz 1998), whilst preferential deadenylation may play a role in regulating Hsp70 (and perhaps, therefore, other stress protein) expression in *Drosophila* (Dellavalle *et al.* 1994). The presence of deadenylated mRNAs would clearly reduce the efficiency of systems utilizing a polydT reverse transcription step. The efficiency of any system also depends on the quality of the starting material. All differential display techniques use mRNA as their target material. However, it is difficult to isolate mRNA that is completely free of ribosomal RNA. Even if polydT primers are used to prime first strand cDNA synthesis, ribosomal RNA is often transcribed to some degree (Clontech PCR-Select cDNA Subtraction kit user manual). It has been shown, at least in the case of SSH, that a high rRNA:mRNA ratio can lead to inefficient subtractive hybridization (Clontech PCR-Select cDNA Subtraction kit user manual), and there is no reason to suppose that it will not do likewise in other SH approaches. Finally, those techniques that utilise a presubtraction amplification step (e.g. RDA) may present a skewed representation since some sequences amplify better than others.

Of course, probably the most important consideration is the temporal factor. It is clear that any given differential display experiment can only interrogate a cell at one point in time. It may well be that a high percentage of the genes showing altered expression at that time are obtained. However, given that disease processes and responses to environmental stimuli involve dynamic cascades of signalling, regulation, production and action, it is clear that all those genes which are switched on/off at different times will not be recovered and, therefore, vital information may well be missed. It is, therefore, imperative to obtain as much information about the model system beforehand as possible, from which a strategy can be derived for targeting specific time points or events that are of particular interest to the investigator. One way of getting round this problem of single time point analysis is to conduct the experiment over a suitable time course which, of course, adds substantially to the amount of work involved.

How sensitive are differential expression technologies?

There has been little published data that addresses the issue of how large the change in expression must be for it to permit isolation of the gene in question with the various differential expression technologies. Although the isolation of genes whose expression is changed as little as 1.5-fold has been reported using SSH (Groenink and Leegwater 1996), it appears that those demonstrating a change in excess of 5-fold are more likely to be picked up. Thus, there is a 'grey zone' in between where small changes could fade in and out of isolation between

experiments and animals. DD, on the other hand, is not subject to this grey zone since, unlike SH approaches, it does not amplify the difference in expression between two samples. Wan *et al.* (1996) reported that differences in expression of twofold or more are detectable using DD.

Resolution and visualization of differential expression products

It seems highly improbable with current technology that a gel system could be developed that is able to resolve all gene species showing altered expression in any given test system (be it SH- or DD-based). Polyacrylamide gel electrophoresis (PAGE) can resolve size differences down to 0.2% (Sambrook *et al.* 1989) and are used as standard in DD experiments. Even so, it is clear that a complex series of gene products such as those seen in a DD will contain unresolvable components. Thus, what appears to be one band in a gel may in fact turn out to be several. Indeed, it has been well documented (Mathieu-Daude *et al.* 1996, Smith *et al.* 1997) that a single band extracted from a DD often represents a composite of heterogeneous products, and the same has been found for SSH displays in this laboratory (Rockett *et al.* 1997). One possible solution was offered by Mathieu-Daude *et al.* (1996), who extracted and reamplified candidate bands from a DD display and used single strand conformation polymorphism (SSCP) analysis to confirm which components represented the truly differentially expressed product.

Many scientists often try to avoid the use of PAGE where possible because it is technically more demanding than agarose gel electrophoresis (AGE). Unfortunately, high resolution agarose gels such as Metaphor (FMC, Lichfield, UK) and AquaPor HR (National Diagnostics, Hessle, UK), whilst easier to prepare and manipulate than PAGE, can only separate DNA sequences which differ in size by around 1.5–2% (15–20 base pairs for a 1Kb fragment). Thus, SSH, RDA or other such products which differ in size by less than this amount are normally not resolvable. However, a simple technique does in fact exist for increasing the resolving power of AGE—the inclusion of HA-red (10-phenyl neutral red-PEG ligand) or HA-yellow (bisbenzamide-PEG ligand) (Hanse Analytik GmbH, Bremen, Germany) in a gel separates identical or closely sized products on base content. Specifically, HA-red and -yellow selectively bind to GC and AT DNA motifs, respectively (Wawer *et al.* 1995, Hanse Analytik 1997, personal communication). Since both HA-stains possess an overall positive charge, they migrate towards the cathode when an electric field is applied. This is in direct opposition to DNA, which is negatively charged and, therefore, migrates towards the anode. Thus, if two DNA clones are identical in size (as perceived on a standard high resolution agarose gel), but differ in AT/GC content, inclusion of a HA-dye in the gel will effectively retard the migration of one of the sequences compared to the other, effectively making it apparently larger and, thus, providing a means of differentiating between the two. The use of HA-red has been shown to resolve sequences with an AT variation of less than 1% (Wawer *et al.* 1995), whilst Hanse Analytik have reported that HA staining is so sensitive that in one case it was used to distinguish two 567bp sequences which differed by only a single point mutation (Hanse Analytik 1996, personal communication). Therefore, if one wishes to check whether all the clones produced from a specific band in a differential display experiment are derived from the same gene species, a small amount of reamplified or digested clone can be run on a standard high resolution gel, and a second aliquot

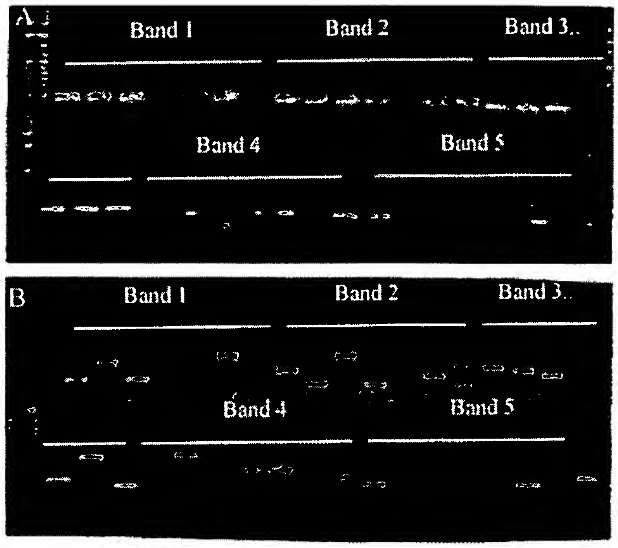


Figure 10. Discrimination of clones of identical/nearly identical size using HA-red. Bands of decreasing size (1–5) were extracted from the final display of a suppression subtractive hybridization experiment and cloned. Seven colonies were picked at random from each cloned band and their inserts amplified using PCR. The products were run on two gels, (A) a high resolution 2% agarose gel, and (B) a high resolution 2% agarose gel containing 1 U/ml HA-red. With few exceptions, all the clones from each band appear to be the same size (gel A). However, the presence of HA-red (gel B), which separates identically-sized DNA fragments based on the percentage of GC within the sequence, clearly indicates the presence of different gene species within each band. For example, even though all five re-amplified clones of band 1 appear to be the same size, at least four different gene species are represented.

in a similar gel containing one of the HA-stains. The standard gel should indicate any gross size differences, whilst the HA-stained gel should separate otherwise unresolvable species (on standard AGE) according to their base content. Geisinger *et al.* (1997) reported successful use of this approach for identifying DD-derived clones. Figure 10 shows such an experiment carried out in this laboratory on clones obtained from a band extracted from an SSH display.

An alternative approach is to carry out a 2-D analysis of the differential display products. In this approach, size-based separation is first carried out in a standard agarose gel. The gel slice containing the display is then extracted and incorporated in to a HA gel for resolution based on AT/GC content.

Of course, one should always consider the possibility of there being different gene species which are the same size and have the same GC/AT content. However, even these species are not unresolvable given some effort—again, one might use SSCP, or perhaps a denaturing gradient gel electrophoresis (DGGE) or temperature gradient field electrophoresis (TGGE) approach to resolve the contents of a band, either directly on the extracted band (Suzuki *et al.* 1991) or on the reamplified product.

The requirement of some differential display techniques to visualize large numbers of products (e.g. DD and GEF) can also present a problem in that, in terms of numbers, the resolution of PAGE rarely exceeds 300–400 bands. One approach to overcoming this might be to use 2-D gels such as those described by Uitterlinden *et al.* (1989) and Hatada *et al.* (1991).

Extraction of differentially expressed bands from a gel can be complex since, in some cases (e.g. DD, GEF), the results are visualized by autoradiographic means, such that precise overlay of the developed film on the gel must occur if the correct band is to be extracted for further analysis. Clearly, a misjudged extraction can account for many man-hours lost. This problem, and that of the use of radioisotopes, has been addressed by several groups. For example, Lohmann *et al.* (1995) demonstrated that silver staining can be used directly to visualize DD bands in horizontal PAGs. An *et al.* (1996) avoided the use of radioisotopes by transferring a small amount (20–30%) of the DNA from their DD to a nylon membrane, and visualizing the bands using chemiluminescent staining before going back to extract the remaining DNA from the gel. Chen and Peck (1996) went one step further and transferred the entire DD to a nylon membrane. The DNA bands were then visualized using a digoxigenin (DIG) system (DIG was attached to the polydT primers used in the differential display procedure). Differentially expressed bands were cut from the membrane and the DNA eluted by washing with PCR buffer prior to reamplification.

One of the advantages of using techniques such as SSH and RDA is that the final display can be run on an agarose gel and the bands visualized with simple ethidium bromide staining. Whilst this approach can provide acceptable results, overstaining with SYBR Green I or SYBR Gold nucleic acid stains (FMC) effectively enhances the intensity and sharpness of the bands. This greatly aids in their precise extraction and often reveals some faint products that may otherwise be overlooked. Whilst differential displays stained with SYBR Green I are better visualized using short wavelength UV (254 nm) rather than medium wavelength (306 nm), the shorter wavelength is much more DNA damaging. In practice, it takes only a few seconds to damage DNA extracted under 254 nm irradiation, effectively preventing reamplification and cloning. The best approach is to overstain with SYBR Green I and extract bands under a medium wavelength UV transillumination.

The possible use of 'microfingerprinting' to reduce complexity

Given the sheer number of gene products and the possible complexity of each band, an alternative approach to rapid characterization may be to use an enhanced analysis of a small section of a differential display—a 'sub-fingerprint' or 'micro-fingerprint'. In this case, one could concentrate on those bands which only appear in a particular chosen size region. Reducing the fingerprint in this way has at least two advantages. One is that it should be possible to use different gel types, concentrations and run times tailored exactly to that region. Currently, one might run products from 100–3000 + bp on the same gel, which leads to compromize in the gel system being used and consequently to suboptimal resolution, both in terms of size and numbers, and can lead to problems in the accurate excision of individual bands. Secondly, it may be possible to enhance resolution by using a 2-D analysis using a HA-stain, as described earlier. In summary, if a range of gene product sizes is carefully chosen to included certain 'relevant' genes, the 2-D system standardized, and appropriate gene analysis used, it may be possible to develop a method for the early and rapid identification of compounds which have similar or widely different cellular effects. If the prognosis for exposure to one or more other chemicals which display a similar profile is already known, then one could perhaps predict similar effects for any new compounds which show a similar micro-fingerprint.

An alternative approach to microfingerprinting is to examine altered expression in specific families of genes through careful selection of PCR primers and/or post-reaction analysis. Stress genes, growth factors and/or their receptors, cell cycling genes, cytochromes P450 and regulatory proteins might be considered as candidates for analysis in this way. Indeed, some off-the-shelf DNA arrays (e.g. Clontech's Atlas cDNA Expression Array series) already anticipated this to some degree by grouping together genes involved in different responses e.g. apoptosis, stress, DNA-damage response etc.

Screening

False positives

The generation of false positives has been discussed at length amongst the differential display community (Liang *et al.* 1993, 1995, Nishio *et al.* 1994, Sun *et al.* 1994, Sompayrac *et al.* 1995). The reason for false positives varies with the technique being used. For instance, in RDA, the use of adaptors which have not been HPLC purified can lead to the production of false positives through illegitimate ligation events (O'Neill and Sinclair 1997), whilst in DD they can arise through PCR artifacts and illegitemate transcription of rRNA. In SH, false positives appear to be derived largely from abundant gene species, although some may arise from cDNA/mRNA species which do not undergo hybridization for technical reasons.

A quick screening of putative differentially expressed clones can be carried out using a simple dot blot approach, in which labelled first strand probes synthesized from tester and driver mRNA are hybridized to an array of said clones (Hedrick *et al.* 1984, Sakaguchi *et al.* 1986). Differentially expressed clones will hybridize to tester probe, but not driver. The disadvantage of this approach is that rare species may not generate detectable hybridization signals. One option for those using SSH is to screen the clones using a labelled probe generated from the subtracted cDNA from which it was derived, and with a probe made from the reverse subtraction reaction (ClonTechniques 1997a). Since the SSH method enriches rare sequences, it should be possible to confirm the presence of clones representing low abundance genes. Despite this quick screening step, there is still the need to go back to the original mRNA and confirm the altered expression using a more quantitative approach. Although this may be achieved using Northern blots, the sensitivity is poor by today's high standards and one must rely on PCR methods for accurate and sensitive determinations (see below).

Sequence analysis

The majority of differential display procedures produce final products which are between 100 and 1000bp in size. However, this may considerably reduce the size of the sequence for analysis of the DNA databases. This in turn leads to a reduced confidence in the result—several families of genes have members whose DNA sequences are almost identical except in a few key stretches, e.g. the cytochrome P450 gene superfamily (Nelson *et al.* 1996). Thus, does the clone identified as being almost identical to gene X_0 really come from that gene, or its brother gene X_1 or its as yet undiscovered sister X_2? For example, using SSH, part of a gene was isolated,

which was up-regulated in the liver of rats exposed to Wy-14,643 and was identified by a FASTA search as being transferrin (data not shown). However, transferrin is known to be downregulated by hypolipidemic peroxisome proliferators such as Wy-14,643 (Hertz *et al.* 1996), and this was confirmed with subsequent RT-PCR analysis. This suggests that the gene sequence isolated may belong to a gene which is closely related to transferrin, but is regulated by a different mechanism.

A further problem associated with SH technology is redundancy. In most cases before SH is carried out, the cDNA population must first be simplified by restriction digestion. This is important for at least two reasons:

(1) To reduce complexity—long cDNA fragments may form complex networks which prevent the formation of appropriate hybrids, especially at the high concentrations required for efficient hybridization.
(2) Cutting the cDNAs into small fragments provides better representation of individual genes. This is because genes derived from related but distinct members of gene families often have similar coding sequences that may cross-hybridize and be eliminated during the subtraction procedure (Ko 1990). Furthermore, different fragments from the same cDNA may differ considerably in terms of hybridization and amplification and, thus, may not efficiently do one or the other (Wang and Brown 1991). Thus, some fragments from differentially expressed cDNAs may be eliminated during subtractive hybridization procedures. However, other fragments may be enriched and isolated. As a consequence of this, some genes will be cut one or more times, giving rise to two or more fragments of different sizes. If those same genes are differentially expressed, then two or more of the different size fragments may come through as separate bands on the final differential display, increasing the observed redundancy and increasing the number of redundant sequencing reactions.

Sequence comparisons also throw up another important point—at what degree of sequence similarity does one accept a result. Is 90% identitiy between a gene derived from your model species and another acceptably close? Is 95% between your sequence and one from the same species also acceptable? This problem is particularly relevant when the forward and reverse sequence comparisons give similar sequences with completely different gene species! An arbitrary decision seems to be to allocate genes that are definite (95% and above similarity) and then group those between 60 and 95% as being related or possible homologues.

Quantitative analysis

At some point, one must give consideration to the quantitative analysis of the candidate genes, either as a means of confirming that they are truly differentially expressed, or in order to establish just what the differences are. Northern blot analysis is a popular approach as it is relatively easy and quick to perform. However, the major drawback with Northern blots is that they are often not sensitive enough to detect rare sequences. Since the majority of messages expressed in a cell are of low abundance (see table 1), this is a major problem. Consequently, RT-PCR may be the method of choice for confirming differential expression. Although the procedure is somewhat more complex than Northern analysis, requiring synthesis of primers and optimization of reaction conditions for each gene species, it is now possible to set up high throughput PCR systems using mulitchannel pipettes, 96 +-well plates and

appropriate thermal cycling technology. Whilst quantitative analysis is more desirable, being more accurate and without reliance on an internal standard, the money and time needed to develop a competitor molecule is often excessive, especially when one might be examining tens or even hundreds of gene species. The use of semi-quantitative analysis is simpler, although still relatively involved. One must first of all choose an internal standard that does not change in the test cells compared to the controls. Numerous reference genes have been tried in the past, for example interferon-gamma (IFN-γ, Frye *et al.* 1989), β-actin (Heuval *et al.* 1994), glyceraldehyde-3-phosphate dehydrogenase (GAPDH, Wong *et al.* 1994), dihydrofolate reductase (DHFR, Mohler and Butler 1991), β-2-microglobulin (β-2-m, Murphy *et al.* 1990), hypoxanthine phosphoribosyl transferase (HPRT, Foss *et al.* 1998) and a number of others (ClonTechniques 1997b). Ideally, an internal standard should not change its level of expression in the cell regardless of cell age, stage in the cell cycle or through the effects of external stimuli. However, it has been shown on numerous occasions that the levels of most housekeeping genes currently used by the research community do in fact change under certain conditions and in different tissues (ClonTechniques 1997b). It is imperative, therefore, that preliminary experiments be carried out on a panel of housekeeping genes to establish their suitability for use in the model system.

Interpretation of quantitative data must also be treated with caution. By comparing the lists of genes identified by differential expression one can perhaps gain insight into why two different species react in different ways to external stimuli. For example, rats and mice appear sensitive to the non-genotoxic effects of a wide range of peroxisome proliferators whilst Syrian hamsters and guinea pigs are largely resistant (Orton *et al.* 1984, Rodricks and Turnbull 1987, Lake *et al.* 1989, 1993, Makowska *et al.* 1992). A simplified approach to resolving the reason(s) why is to compare lists of up- and down-regulated genes in order to identify those which are expressed in only one species and, through background knowledge of the effects of the said gene, might suggest a mechanism of facilitated non-genotoxic carcinogenesis or protection. Of course, the situation is likely to be far more complex. Perhaps if there were one key gene protecting guinea pig from non-genotoxic effects and it was upregulated 50 times by PPs, the same gene might only be up-regulated five times in the rat. However, since both were noted to be upregulated, the importance of the gene may be overlooked. Just to complicate matters, a large change in expression does not necessarily mean a biologically important change. For example, what is the true relevance of gene Y which shows a 50-fold increase after a particular treatment, and gene Z which shows only a 5-fold increase? If one examines the literature one may find that historically, gene Y has often been shown to be up-regulated 40–60-fold by a number of unrelated stimuli—in light of this the 50-fold increase would appear less significant. However, the literature may show that gene Z has never been recorded as having more than doubled in expression—which makes your 5-fold increase all the more exciting. Perhaps even more interesting is if that same 5-fold increase has only been seen in related neoplasms or following treatment with related chemicals.

Problems in using the differential display approach

Differential display technology originally held promise of an easily obtainable 'fingerprint' of those genes which are up- or down-regulated in test animals/cells in a developmental process or following exposure to given stimuli. However, it has

become clear that the fingerprinting process, whilst still valid, is much too complex to be represented by a single technique profile. This is because all differential display techniques have common and/or unique technical problems which preclude the isolation and identification of all those genes which show changes in expression. Furthermore, there are important genetic changes related to disease development which differential expression analysis is simply not designed to address. An example of this is the presence of small deletions, insertions, or point mutations such as those seen in activated oncogenes, tumour suppressor genes and individual polymorphisms. Polymorphic variations, small though they usually are, are often regarded as being of paramount importance in explaining why some patients respond better than others to certain drug treatments (and, in logical extension, why some people are less affected by potentially dangerous xenobiotics/carcinogens than others). The identification of such point mutations and naturally occurring polymorphisms requires the subsequent application of sequencing, SSCP, DGGE or TGGE to the gene of interest. Furthermore, differential display is not designed to address issues such as alternatively spliced gene species or whether an increased abundance of mRNA is a result of increased transcription or increased mRNA stability.

Conclusions

Perhaps the main advantage of open system differential display techniques is that they are not limited by extant theories or researcher bias in revealing genes which are differentially expressed, since they are designed to amplify all genes which demonstrate altered expression. This means that they are useful for the isolation of previously unknown genes which may turn out be useful biomarkers of a particular state or condition. At least one open system (SAGE) is also quantitative, thus eliminating the need to return to the original mRNA and carry out Northern/PCR analysis to confirm the result. However, the rapid progress of genome mapping projects means that over the next 5–10 years or so, the balance of experimental use will switch from open to closed differential display systems, particularly DNA arrays. Arrays are easier and faster to prepare and use, provide quantitative data, are suitable for high throughput analysis and can be tailored to look at specific signalling pathways or families of genes. Identification of all the gene sequences in human and common laboratory animals combined with improved DNA array technology, means that it will soon no longer be necessary to try to isolate differentially expressed genes using the technically more demanding open system approach. Thus, their main advantage (that of identifying unknown genes) will be largely eradicated. It is likely, therefore, that their sphere of application will be reduced to analysis of the less common laboratory species, since it will be some time yet before the genomes of such animals as zebrafish, electric eels, gerbils, crayfish and squid, for example, will be sequenced.

Of course, in the end the question will always remain: What is the functional/biological significance of the identified, differentially expressed genes? One persistent problem is understanding whether differentially expressed genes are a cause or consequence of the altered state. Furthermore, many chemicals, such as non-genotoxic carcinogens, are also mitogens and so genes associated with replication will also be upregulated but may have little or nothing to do with the

carcinogenic effect. Whilst differential display technology cannot hope to answer these questions, it does provide a springboard from which identification, regulatory and functional studies can be launched. Understanding the molecular mechanism of cellular responses is almost impossible without knowing the regulation and function of those genes and their condition (e.g. mutated). In an abstract sense, differential display can be likened to a still photograph, showing details of a fixed moment in time. Consider the Historian who knows the outcome of a battle and the placement and condition of the troops before the battle commenced, but is asked to try and deduce how the battle progressed and why it ended as it did from a few still photographs—an impossible task. In order to understand the battle, the Historian must find out the capabilities and motivation of the soldiers and their commanding officers, what the orders were and whether they were obeyed. He must examine the terrain, the remains of the battle and consider the effects the prevailing weather conditions exerted. Likewise, if mechanistic answers are to be forthcoming, the scientist must use differential display in combination with other techniques, such as knockout technology, the analysis of cell signalling pathways, mutation analysis and time and dose response analyses. Although this review has emphasized the importance of differential gene profiling, it should not be considered in isolation and the full impact of this approach will be strengthened if used in combination with functional genomics and proteomics (2-dimensional protein gels from isoelectric focusing and subsequent SDS electrophoresis and virtual 2D-maps using capillary electrophoresis). Proteomics is attracting much recent attention as many of the changes resulting in differential gene expression do not involve changes in mRNA levels, as decribed extensively herein, but rather protein–protein, protein–DNA and protein phosphorylation events which would require functional genomics or proteomic technologies for investigation.

Despite the limitations of differential display technology, it is clear that many potential applications and benefits can be obtained from characterizing the genetic changes that occur in a cell during normal and disease development and in response to chemical or biological insult. In light of functional data, such profiling will provide a 'fingerprint' of each stage of development or response, and in the long term should help in the elucidation of specific and sensitive biomarkers for different types of chemical/biological exposure and disease states. The potential medical and therapeutic benefits of understanding such molecular changes are almost immeasurable. Amongst other things, such fingerprints could indicate the family or even specific type of chemical an individual has been exposed to plus the length and/or acuteness of that exposure, thus indicating the most prudent treatment. They may also help uncover differences in histologically identical cancers, provide diagnostic tests for the earliest stages of neoplasia and, again, perhaps indicate the most efficacious treatment.

The Human Genome Project will be completed early in the next century and the DNA sequence of all the human genes will be known. The continuing development and evolution of differential gene expression technology will ensure that this knowledge contributes fully to the understanding of human disease processes.

Acknowledgements

We acknowledge Drs Nick Plant (University of Surrey), Sally Darney and Chris Luft (US EPA at RTP) for their critical analysis of the manuscript prior to submission. This manuscript has been reviewed in accordance with the policy of the

US Environmental Protection Agency and approved for publication. Approval does not signify that the contents reflect the views and policies of the Agency, nor does mention of trade names constitute endorsement or recommendation for use.

References

ADAMS, M. D., KELLEY, J. M., GOCAYNE, J. D., DUBNICK, M., POLYMEROPOULOS, M. H., XIAO, H., MERRIL, C. R., WU, A., OLDE, B., MORENO, R. F., KERLAVAGE, A. R., MCCOMBIE, W. R. and VENTOR, J. C., 1991, Complementary DNA sequencing: expressed sequence tags and human genome project. *Science*, **252**, 1651–1656.

AN, G., LUO, G., VELTRI, R. W. and O'HARA, S. M., 1996, Sensitive non-radioactive differential display method using chemiluminescent detection. *Biotechniques*, **20**, 342–346.

AXEL, R., FEIGELSON, P. and SCHULTZ, G., 1976, Analysis of the complexity and diversity of mRNA from chicken liver and oviduct. *Cell*, **7**, 247–254.

BAND, V. and SAGER, R., 1989, Distinctive traits of normal and tumor-derived human mammary epithelial cells expressed in a medium that supports long-term growth of both cell types. *Proceedings of the Naional Academy of Sciences, USA*, **86**,1249–1253.

BAUER, D., MULLER, H., REICH, J., RIEDEL, H., AHRENKIEL, V., WARTHOE, P. and STRAUSS, M., 1993, Identification of differentially expressed mRNA species by an improved display technique (DDRT-PCR). *Nucleic Acids Research*, **21**, 4272–4280.

BERTIOLI, D. J., SCHLICHTER, U. H. A., ADAMS, M. J., BURROWS, P. R., STEINBISS, H.-H. and ANTONIW, J. F., 1995, An analysis of differential display shows a strong bias towards high copy number mRNAs. *Nucleic Acids Research*, **23**, 4520–4523.

BRAVO, R., 1990, Genes induced during the G0/G1 transition in mouse fibroblasts. *Seminars in Cancer Biology*, **1**, 37–46.

BURN, T. C., PETROVICK, M. S., HOHAUS, S., ROLLINS, B. J. and TENEN, D. G., 1994, Monocyte chemoattractant protein-1 gene is expressed in activated neutrophils and retinoic acid-induced human myeloid cell lines. *Blood*, **84**, 2776–2783.

CAO, J., CAI, X., ZHENG, L., GENG, L., SHI, Z., PAO, C. C. and ZHENG, S., 1997, Characterisation of colorectal cancer-related cDNA clones obtained by subtractive hybridisation screening. *Journal of Cancer Research and Clinical Oncology*, **123**, 447–451.

CASSIDY, S. B., 1995, Uniparental disomy and genomic imprinting as causes of human genetic disease. *Environmental and Molecular Mutagenesis*, **25** (Suppl 26), 13–20.

CHANG, G. W. and TERZAGHI-HOWE, M., 1998, Multiple changes in gene expression are associated with normal cell-induced modulation of the neoplastic phenotype. *Cancer Research*, **58**, 4445–4452.

CHEN, J., SCHWARTZ, D. A., YOUNG, T. A., NORRIS, J. S. and YAGER, J. D., 1996, Identification of genes whose expression is altered during mitosuppression in livers of ethinyl estradiol-treated female rats. *Carcinogenesis*, **17**, 2783–2786.

CHEN, J. J. W. and PECK, K., 1996, Non-radioactive differential display method to directly visualise and amplify differential bands on nylon membrane. *Nucleic Acid Research*, **24**, 793–794.

CLON TECHNIQUES, 1997a, PCR-Select Differential Screening Kit—the nextstep after Clontech PCR-Select cDNA subtraction. *ClonTechniques*, **XII**, 18–19.

CLON TECHNIQUES, 1997b, Housekeeping RT-PCR amplimers and cDNA probes. *ClonTechniques*, **XII**, 15–16.

DAVIS, M. M., COHEN, D. I., NIELSEN, E. A., STEINMETZ, M., PAUL, W. E. and HOOD, L., 1984, Cell-type-specific cDNA probes and the murine I region: the localization and orientation of Ad alpha. *Proceedings of the National Academy of Sciences (USA)*, **81**, 2194–2198.

DELLAVALLE, R. P., PETERSON, R. and LINDQUIST, S., 1994, Preferential deadenylation of HSP70 mRNA plays a key role in regulating Hsp70 expression in Drosophila melanogaster. *Molecular and Cell Biology*, **14**, 3646–3659.

DERISI, J. L., VASHWANATH, R. L. and BROWN, P., 1997, Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science*, **278**, 680–686.

DIATCHENKO, L., LAU, Y.-F. C., CAMPBELL, A. P., CHENCHIK, A., MOQADAM, F., HUANG, B., LUKYANOV, K., GURSKAYA, N., SVERDLOV, E. D. and SIEBERT, P. D., 1996, Suppression subtractive hybridisation: A method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proceedings of the National Academy of Sciences (USA)*, **93**, 6025–6030.

DOGRA, S. C., WHITELAW, M. L. and MAY, B. K., 1998, Transcriptional activation of cytochrome P450 genes by different classes of chemical inducers. *Clinical and Experimental Pharmacology and Physiology*, **25**, 1–9.

DUGUID, J. R. and DINAUER, M. C., 1990, Library subtraction of *in vitro* cDNA libraries to identify differentially expressed genes in scrapie infection. *Nucleic Acids Research*, **18**, 2789–2792.

DUNBAR, P. R., OGG, G. S., CHEN, J., RUST, N., VAN DER BRUGGEN, P. and CERUNDOLO, V., 1998, Direct isolation, phenotyping and cloning of low-frequency antigen-specific cytotoxic T lymphocytes from peripheral blood. *Current Biology*, **26**, 413–416.

FITZPATRICK ,D. R., GERMAIN -LEE, E. and VALLE, D., 1995, Isolation and characterisation of rat and human cDNAs encoding a novel putative peroxisomal enoyl-CoA hydratase. *Genomics*, **27**, 457–466.

FOSS, D. L., BAARSCH, M. J. and MURTAUGH, M. P., 1998, Regulation of hypoxanthine phosphoribosyltransferase, glyceraldehyde-3-phosphate dehydrogenase and beta-actin mRNA expression in porcine immune cells and tissues. *Animal Biotechnology*, **9**, 67–78 .

FRYE, R. A., BENZ, C. C. and LIU, E., 1989, Detection of amplified oncogenes by differential polymerase chain reaction. *Oncogene*, **4**, 1153–1157.

GEISINGER , A., RODRIGUEZ, R., ROMERO , V. and WETTSTEIN R., 1997, A simple method for screening cDNAs arising from the cloning of RNA differential display bands. *Elsevier Trends Journals Technical Tips Online*, http://tto.trends.com, document T01110.

GRESS, T. M., HOHEISEL , J. D., LENNON , G. G., ZEHETNER , G. and LEHRACH , H., 1992, Hybridisation fingerprinting of high density cDNA filter arrays with cDNA pools derived from whole tissues. *Mammalian Genome*, **3**, 609–619.

GRIFFIN , G. and KRISHNA , S., 1998, Cytokines in infectious diseases. *Journal of the Royal College of Physicians, London*, **32**, 195–198.

GROENINK , M. and LEEGWATER , A. C. J., 1996, Isolation of delayed early genes associated with liver regeneration using Clontech PCR-select subtraction technique. *Clontechniques*, **XI**, 23–24.

GUIMARAES , M. J., BAZAN, J. F., ZLOTNIK , A., WILES , M. V., GRIMALDI , J. C., LEE, F. and MCCLANAHAN , T., 1995b, A new approach to the study of haematopoietic development in the yolk sac and embryoid bodies. *Development*, 121, 3335–3346.

GUIMERAES , M. J., LEE, F., ZLOTNIK , A. and MCCLANAHAN , T., 1995a, Differential display by PCR:novel findings and applications. *Nucleic Acids Research*, **23**, 1832–1833.

GURSKAYA, N. G., DIATCHENKO , L., CHENCHIK , P. D., SIEBERT , P. D., KHASPEKOV , G. L., LUKYANOV , K. A., VAGNER, L. L., ERMOLAEVA , O. D., LUKYANOV, S. A. and SVERDLOV , E. D., 1996, Equalising cDNA subtraction based on selective suppression of polymerase chain reaction: Cloning of Jurkat cell transcripts induced by phytohemaglutinin and phorbol 12-Myrystate 13-Acetate. *Analytical Biochemistry*, **240**, 90–97.

HAMPSON , I. N. and HAMPSON , L., 1997, CCLS and DROP—subtractive cloning made easy. *Life Science News* (A publication of Amersham Life Science), **23**, 22–24.

HAMPSON , I. N., HAMPSON , L. and DEXTER , T. M., 1996, Directional random oligonucleotide primed (DROP) global amplification of cDNA: its application to subtractive cDNA cloning. *Nucleic Acids Research*. **24**, 4832–4835.

HAMPSON , I. N., POPE, L., COWLING , G. J. and DEXTER , T. M., 1992, Chemical cross linking subtraction (CCLS): a new method for the generation of subtractive hybridisation probes. *Nucleic Acids Research*, **20**, 2899.

HARA, E., KATO, T., NAKADA, S., SEKIYA , S. and ODA, K., 1991, Subtractive cDNA cloning using oligo(dT)30-latex and PCR: isolation of cDNA clones specific to undifferentiated human embryonal carcinoma cells. *Nucleic Acids Research*, **19**, 7097–7104.

HATADA, I., HAYASHIZAKE , Y., HIROTSUNE , S., KOMATSUBARA , H. and MUKAI, T., 1991, A genomic scanning method for higher organisms using restriction sites as landmarks. *Proceedings of the National Academy of Sciences (USA)*, **88**, 9523–9527.

HECHT , N., 1998, Molecular mechanisms of male sperm cell differentiation. *Bioessays*, **20**, 555–561.

HEDRICK, S., COHEN , D. I., NIELSEN , E. A. and DAVIS , M. E., 1984, Isolation of T cell-specific membrane-associated proteins. *Nature*, **308**, 149–153.

HERTZ, R., SECKBACH , M., ZAKIN , M. M. and BAR-TANA, J., 1996, Transcriptional suppression of the transferrin gene by hypolipidemic peroxisome proliferators. *Journal of Biological Chemistry*, **271**, 218–224.

HEUVAL , J. P. V., CLARK, G. C., KOHN , M. C., TRITSCHER , A. M., GREENLEE , W. F., LUCIER , G. W. and BELL , D. A., 1994, Dioxin-responsive genes: Examination of dose-response relationships using quantitative reverse transciptase-polymerase chain reaction. *Cancer Research*, **54**, 62–68.

HILLIER , L. D., LENNON , G., BECKER , M., BONALDO , M. F., CHIAPELLI , B., CHISSOE , S., DIETRICH , N., DUBUQUE, T., FAVELLO , A., GISH , W., HAWKINS , M., HULTMAN , M., KUCABA, T., LACY, M., LE, M., LE, N., MARDIS, E., MOORE, B., MORRIS, M., PARSONS, J., PRANGE, C., RIFKIN , L., ROHLFING , T., SCHELLENBERG , K., SOARES, M. B., TAN, F., THIERRY -MEG, J., TREVASKIS , E., UNDERWOOD , K., WOHLDMAN , P., WATERSTON , R., WILSON , R and MARRA, M., 1996, Generation and analysis of 280,000 human expressed sequence tags. *Genome Research*, **6**, 807–828.

HUBANK, M. and SCHATZ, D. G., 1994, Identifying differences in mRNA expression by representational difference analysis. *Nucleic Acids Research*, **22**, 5640–5648.

HUNTER , T., 1991, Cooperation between oncogenes. *Cell*, **64**, 249–270.

IVANOVA , N. B. and BELYAVSKY , A. V., 1995, Identification of differentially expressed genes by restriction endonuclease-based gene expression fingerprinting. *Nucleic Acids Research*, **23**, 2954–2958.

JAMES , B. D. and HIGGINS , S. J, 1985, *Nucleic Acid Hybridisation* (Oxford: IRL Press Ltd).

KAS-DEELEN , A. M., HARMSEN , M. C., DE MAAR, E. F. and VAN SON, W. J, 1998, A sensitive method for

quantifying cytomegalic endothelial cells in peripheral blood from cytomegalovirus-infected patients. *Clinical Diagnostic and Laboratory Immunology*, **5**, 622–626.

KILTY, I. and VICKERS, P., 1997, Fractionating DNA fragments generated by differential display PCR. *Strategies Newsletter* (Stratagene), **10**, 50–51.

KLEINJAN, D.-J. and VAN HEYNINGEN, V., 1998, Position effect in human genetic disease. *Human and Molecular Genetics*, **7**, 1611–1618.

KO, M. S., 1990, An 'equalized cDNA library' by the reassociation of short double-stranded cDNAs. *Nucleic Acids Research*, **18**, 5705–5711.

LAKE, B. G., EVANS, J. G., CUNNINGHAME, M. E. and PRICE, R. J., 1993, Comparison of the hepatic effects of Wy-14,643 on peroxisome proliferation and cell replication in the rat and Syrian hamster. *Environmental Health Perspectives*, **101**, 241–248.

LAKE, B. G., EVANS, J. G., GRAY, T. J. B., KOROSI, S. A. and NORTH, C. J., 1989, Comparative studies of nafenopin-induced hepatic peroxisome proliferation in the rat, Syrian hamster, guiea pig and marmoset. *Toxicology and Applied Pharmacology*, **99**, 148–160.

LENNARD, M. S., 1993, Genetically determined adverse drug reactions involving metabolism. *Drug Safety*, **9**, 60–77.

LEVY, S., TODD, S. C. and MAECKER, H. T., 1998, CD81(TAPA-1): a molecule involved in signal transduction and cell adhesion in the immune system. *Annual Review of Immunology*, **16**, 89–109.

LIANG, P. and PARDEE, A. B., 1992, Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science*, **257**, 967–971.

LIANG, P., AVERBOUKH, L., KEYOMARSI, K., SAGER, R. and PARDEE, A., 1992, Differential display and cloning of messenger RNAs from human breast cancer versus mammary epithelial cells. *Cancer Research*, **52**, 6966–6968.

LIANG, P., AVERBOUKH, L. and PARDEE, A. B., 1993, Distribution & cloning of eukaryotic mRNAs by means of differential display refinements and optimisation. *Nucleic Acids Research*, **21**, 3269–3275.

LIANG, P., BAUER, D., AVERBOUKH, L., WARTHOE, P., ROHRWILD, M., MULLER, H., STRAUSS, M. and PARDEE, A. B., 1995, Analysis of altered gene expression by differential display. *Methods in Enzymology*, **254**, 304–321.

LINSKENS, M. H., FENG, J., ANDREWS, W. H., ENLOW, B. E., SAATI, S. M., TONKIN, L. A., FUNK, W. D. and VILLEPONTEAU, B., 1995, Cataloging altered gene expression in young and senescent cells using enhanced differential display. *Nucleic Acids Research*, **23**, 3244–3251.

LISITSYN, N., LISIITSYN, N. and WIGLER, M., 1993, Cloning the differences between two complex genomes. *Science*, **259**, 946–951.

LOHMANN, J., SCHICKLE, H. and BOSCH, T. C. G., 1995, REN Display, a rapid and efficient method for non-radioactive differential display and mRNA isolation. *Biotechniques*, **18**, 200–202.

LUNNEY, J. K., 1998, Cytokines orchestrating the immune response. *Reviews in Science and Techology*, **17**, 84–94.

MAKOWSKA, J. M., GIBSON, G. G. and BONNER, F. W., 1992, Species differences in ciprofibrate-induction of hepaic cytochrome P4504A1 and peroxisome proliferation. *Journal of Biochemical Toxicology*, **7**, 183–191.

MALDARELLI, F., XIANG, C., CHAMOUN, G. and ZEICHNER, S. L., 1998, The expression of the essential nuclear splicing factor SC35 is altered by human immunodeficiency virus infection. *Virus Research*, **53**, 39–51.

MATHIEU-DAUDE, F., CHENG, R., WELSH, J. and MCCLELLAND, M., 1996, Screening of differentially amplified cDNA products from RNA arbitrarily primed PCR fingerprints using single strand conformation polymorphism (SSCP) gels. *Nucleic Acids Research*, **24**, 1504–1507.

MCKENZIE, D. and DRAKE, D., 1997, Identification of differentially expressed gene products with the castaway system. *Strategies Newsletter* (Stratagene), **10**,19–20.

MCCLELLAND, M., MATHIEU-DAUDE, F. and WELSH, J., 1996, RNA fingerprinting and differential display using arbitrarily primed PCR. *Trends in Genetics*, **11**, 242–246.

MECHLER, B. and RABBITTS, T. H., 1981, Membrane-bound ribosomes of myeloma cells. IV. mRNA complexity of free and membrane-bound polysomes. *Journal of Cell Biology*, **88**, 29–36.

MEYER, U. A. and ZANGER, U. M., 1997, Molecular mechanisms of genetic polymorphisms of drug metabolism. *Annual Review of Pharmacology and Toxicology*, **37**, 269–296.

MOHLER, K. M. and BUTLER, L. D., 1991, Quantitation of cytokine mRNA levels utilizing the reverse transcriptase-polymerase chain reaction following primary antigen-specific sensitization in vivo—I. Verification of linearity, reproducibility and specificity. *Molecular Immunology*, **28**, 437–447.

MURPHY, L. D., HERZOG, C. E., RUDICK, J. B., TITO FOJO, A. and BATES, S. E., 1990, Use of the polymerase chain reaction in the quantitation of the mdr-1 gene expression. *Biochemistry*, **29**, 10351–10356.

NELSON, D. R., KOYMANS, L., KAMATAKI, T., STEGEMAN, J. J., FEYEREISEN, R., WAXMAN, D. J., WATERMAN, M. R., GOTOH, O., COON, M. J., ESTABTROOK, R. W., GUNSALUS, I. C. and NEBERT, D. W., 1996, Update on new sequences, gene mapping, accession numbers and nomenclature. *Pharmacogenetics*, **6**, 1–42.

NISHIO, Y., AIELLO, L. P. and KING, G. L., 1994, Glucose induced genes in bovine aortic smooth muscle cells identified by mRNA differential display. *FASEB Journal*, **8**, 103–106.

O'NEILL, M. J. and SINCLAIR, A. H., 1997, Isolation of rare transcripts by representational difference analysis. *Nucleic Acids Research*, **25**, 2681–2682.

ORTON, T. C., ADAM, H. K., BENTLEY, M., HOLLOWAY, B. and TUCKER, M. J., 1984, Clobuzarit: species differences in the morphological and biochemical response of the liver following chronic administration. *Toxicology and Applied Pharmacology*, **73**, 138–151.

PELKONEN, O., MAENPAA, J., TAAVITSAINEN, P., RAUTIO, A. and RAUNIO, H., 1998, Inhibition and Induction of human cytochrome P450 (CYP) enzymes. *Xenobiotica*, **28**, 1203–1253.

PHILIPS, S. M., BENDALL, A. J. and RAMSHAW, I. A., 1990, Isolation of genes associated with high metastatic potential in rat mammary adenocarcinomas. *Journal of the National Cancer Institute*, **82**, 199–203.

PRASHAR, Y. and WEISSMAN, S. M., 1996, Analysis of differential gene expression by display of 3'end restriction fragments of cDNAs. *Proceedings of the National Academy of Sciences (USA)*, **93**, 659–663.

RAGNO, S., ESTRADA, I., BUTLER, R. and COLSTON, M. J., 1997, Regulation of macrophage gene expression following invasion by *Mycobacterium tuberculosis*. *Immunology Letters*, **57**, 143–146.

RAMANA, K. V. and KOHLI, K. K., 1998, Gene regulation of cytochrome P450—an overview. *Indian Journal of Experimental Biology*, **36**, 437–446.

RICHARD, L., VELASCO, P. and DETMAR, M., 1998, A simple immunomagnetic protocol for the selective isolation and long-term culture of human dermal microvascular endothelial cells. *Experimental Cell Research*, **240**, 1–6.

ROCKETT, J. C., ESDAILE, D. J. and GIBSON, G. G., 1997, Molecular profiling of non-genotoxic hepatocarcinogenesis using differential display reverse transcription-polymerase chain reaction (ddRT-PCR). *European Journal of Drug. Metabolism and Pharmacokinetics*, **22**, 329–333.

RODRICKS, J. V. and TURNBULL, D., 1987, Inter-species differences in peroxisomes and peroxisome proliferation. *Toxicology and Industrial Health*, **3**, 197–212.

ROGLER, G., HAUSMANN, M., VOGL, D., ASCHENBRENNER, E., ANDUS, T., FALK, W., ANDREESEN, R., SCHOLMERICH, J. and GROSS, V., 1998, Isolation and phenotypic characterization of colonic macrophages. *Clinical and Experimental Immunology*, **112**, 205–215.

ROHN, W. M., LEE, Y. J. and BENVENISTE, E. N., 1996, Regulation of class II MHC expression. *Critical Reviews in Immunology*, **16**, 311–330.

RUDIN, C. M. and THOMPSON, C. B., 1998, B-cell development and maturation. *Seminars in Oncology*, **25**, 435–446.

SAKAGUCHI, N., BERGER, C. N. and MELCHERS, F., 1986, Isolation of a cDNA copy of an RNA species expressed in murine pre-B cells. *EMBO Journal*, **5**, 2139–2147.

SAMBROOK, J., FRITSCH, E. F. and MANIATIS, T., 1989, Gel electrophoresis of DNA. In N. Ford, M. Nolan and M. Fergusen (eds), *Molecular Cloning—A laboratory manual*, 2nd edition (New York: Cold Spring Harbour Laboratory Press), Volume 1, pp. 6–37.

SARGENT, T. D. and DAWID, I. B., 1983, Differential gene expression in the gastrula of Xenopus laevis. *Science*, **222**, 135–139.

SCHENA, M., SHALON, D., HELLER, R., CHAI, A., BROWN., P. O. and DAVIS, R. W., 1996, Parallel human genome analysis: Microarray-based expression monitoring of 1000 genes. *Proceedings of the National Academy of Sciences (USA)*, **93**, 10614–10619.

SCHNEIDER, C., KING, R. M. and PHILIPSON, L., 1988, Genes specifically expressed at growth arrest of mammalian cells. *Cell*, **54**, 787–793.

SCHNEIDER-MAUNOURY, S., GILARDI-HEBENSTREIT, P. and CHARNAY, P., 1998, How to build a vertebrate hindbrain. Lessons from genetics. *C R Academy of Science III*, **321**, 819–834.

SEMENZA, G. L., 1994, Transcriptional regulation of gene expression: mechanisms and pathophysiology. *Human Mutations*, **3**, 180–199.

SEWALL, C. H., BELL, D. A., CLARK, G. C., TRITSCHER, A. M., TULLY, D. B., VANDEN HEUVEL, J. and LUCIER, G. W., 1995, Induced gene transcription: implications for biomarkers. *Clinical Chemistry*, **41**, 1829–1834.

SINGH, N., AGRAWAL, S. and RASTOGI, A. K., 1997, Infectious diseases and immunity: special reference to major histocompatibility complex. *Emerging Infectious Diseases*, **3**, 41–49.

SMITH, N. R., LI, A., ALDERSLEY, M., HIGH, A. S., MARKHAM, A. F. and ROBINSON, P. A., 1997, Rapid determination of the complexity of cDNA bands extracted from DDRT-PCR polyacrylamide gels. *Nucleic Acids* Research, **25**, 3552–3554.

SOMPAYRAC, L., JANE, S., BURN., T. C., TENEN, D. G. and DANNA, K. J., 1995, Overcoming limitations of the mRNA differential display technique. *Nucleic Acids Research*, **23**, 4738–4739.

ST JOHN, T. P. and DAVIS, R. W., 1979, Isolation of galactose-inducible DNA sequences from Saccharomyces cerevisiae by differential plaque filter hybridisation. *Cell*, **16**, 443–452.

SUN, Y., HEGAMYER, G. and COLBURN, N. H., 1994, Molecular cloning of five messenger RNAs differentially expressed in preneoplastic or neoplastic JB6 mouse epidermal cells: one is homologous to human tissue inhibitor of metalloproteinases-3. *Cancer Research*, **54**, 1139–1144.

SUNG, Y. J. and DENMAN, R. B., 1997, Use of two reverse transcriptases eliminates false-positive results in differential display. *Biotechniques*, **23**, 462–464.

SUTTON, G., WHITE, O., ADAMS, M. and KERLAVAGE, A., 1995, TIGR Assembler; A new tool for assembling large shotgun sequencing projects. *Genome Science and Technology*, **1**, 9–19.

SUZUKI, Y., SEKIYA, T. and HAYASHI, K., 1991, Allele-specific polymerase chain reaction: a method for amplification and sequence determination of a single component among a mixture of sequence variants. *Analytical Biochemistry*, **192**, 82–84.

SYED, V., GU, W. and HECHT, N. B., 1997, Sertoli cells in culture and mRNA differential display provide a sensitive early warning assay system to detect changes induced by xenobiotics. *Journal of Andrology*, **18**, 264–273.

UITTERLINDEN, A. G., SLAGBOOM, P., KNOOK, D. L. and VIJGL, J., 1989, Two-dimensional DNA fingerprinting of human individuals. *Proceedings of the National Academy of Sciences (USA)*, **86**, 2742–2746.

ULLMAN, K. S., NORTHROP, J. P., VERWEIJ, C. L. and CRABTREE, G. R., 1990, Transmission of signals from the T lymphocyte antigen receptor to the genes responsible for cell proliferation and immune function: the missing link. *Annual Review of Immunology*, **8**, 421–452.

VASMATZIS, G., ESSAND, M., BRINKMANN, U., LEE, B. and PASTON, I., 1998, Discovery of three genes specifically expressed in human prostate by expressed sequence tag database analysis. *Proceedings of the National Academy of Sciences (USA)*, **95**, 300-304.

VELCULESCU, V. E., ZHANG, L., VOGELSTEIN, B. and KINZLER, K. W., 1995, Serial analysis of gene expression. *Science*, **270**, 484–487.

VOELTZ, G. K. and STEITZ, J. A., 1998, AuuuA sequences direct mRNA deadenylation uncoupled from decay during Xenopus early development. *Molecular and Cell Biology*, **18**, 7537–7545.

VOGELSTEIN, B. and KINZLER, K. W., 1993, The multistep nature of cancer. *Trends in Genetics*, **9**, 138–141.

WALTER, J., BELFIELD, M., HAMPSON, I. and READ, C., 1997, A novel approach for generating subtractive probes for differential screening by CCLS. *Life Science News*, **21**, 13–14.

WAN, J. S., SHARP, S. J., POIRIER, G. M.-C., WAGAMAN, P. C., CHAMBERS, J., PYATI, J., HOM, Y.-L., GALINDO, J. E., HUVAR, A., PETERSON, P. A., JACKSON, M. R. and ERLANDER, M. G., 1996, Cloning differentially expressed mRNAs. *Nature Biotechnology*, **14**, 1685–1691.

WALTER, J., BELFIELD, M., HAMPSON, I. and READ, C., 1997, A novel approach for generating subtractive probes for differential screening by CCLS, *Life Science News*, **21**, 13–14.

WANG, Z. and BROWN, D. D. 1991, A gene expression screen. *Proceedings of the National Academy of Sciences (USA)*, **88**, 11505–11509.

WAWER, C., RUGGEBERG, H., MEYER, G. and MUYZER, G., 1995, A simple and rapid electrophoresis method to detect sequence variation in PCR-amplified DNA fragments. *Nucleic Acids Research*, **23**, 4928–4929.

WELSH, J., CHADA, K., DALAL, S. S., CHENG, R., RALPH, D. and MCCLELLAND, M., 1992, Arbitrarily primed PCR fingerprinting of RNA. *Nucleic Acids Research*, **20**, 4965–4970.

WONG, H., ANDERSON, W. D., CHENG, T. and RIABOWOL, K. T., 1994, Monitoring mRNA expression by polymerase chain reaction: the 'primer-dropping' method. *Analytical Biochemistry*, **223**, 251–258.

WONG, K. K. and MCCLELLAND, M., 1994, Stress-inducible gene of Salmonella typhimurium identified by arbitrarily primed PCR of RNA. *Proceedings of the National Academy of Sciences (USA)*, **91**, 639-643.

WYNFORD-THOMAS, D., 1991, Oncogenes and anti-oncogenes; the molecular basis of tumour behaviour. *Journal of Pathology*, **165**, 187–201.

XHU, D., CHAN, W. L., LEUNG, B. P., HUANG, F. P., WHEELER, R., PIEDRAFITA, D., ROBINSON, J. H. and LIEW, F. Y., 1998, Selective expression of a stable cell surface molecule on type 2 but not type 1 helper T cells. *Journal of Experimental Medicine*, **187**, 787–794 .

YANG, M. and SYTOWSKI, A. J., 1996, Cloning differentially expressed genes by linker capture subtraction. *Analytical Biochemistry*, **237**, 109–114.

ZHAO, N., HASHIDA, H., TAKAHASHI, N., MISUMI, Y. and SAKAKI, Y., 1995, High-density cDNA filter analysis: a novel approach for large scale quantitative analysis of gene expression. *Gene*, **156**, 207–213.

ZHAO, X. J., NEWSOME, J. T. and CIHLAR, R. L., 1998, Up-regulation of two *candida albicans* genes in the rat model of oral candidiasis detected by differential display. *Microbial Pathogenesis*, **25**, 121–129.

ZIMMERMANN, C. R., ORR, W. C., LECLERC, R. F., BARNARD, C. and TIMBERLAKE, W. E., 1980, Molecular cloning and selection of genes regulated in *Aspergillus* development. *Cell*, **21**, 709–715.

EUROPEAN JOURNAL OF 1997, Vol. 22, No. 4, pp. 329–333

Molecular profiling of non-genotoxic hepatocarcinogenesis using differential display reverse transcription-polymerase chain reaction (ddRT-PCR)

J.C. ROCKETT[1], D.J. ESDAILE[2] and G.G. GIBSON[1]

[1]*Molecular Toxicology Group, School of Biological Sciences, University of Surrey, Guildford, UK*
[2]*Rhône-Poulenc Agrochemicals, Sophia Antipolis, France*

Keywords : ddRT-PCR, non-genotoxic hepatocarcinogenesis, phenobarbital, rat, WY-14,643

SUMMARY

The technique of differential display reverse transcription-polymerase chain reaction (ddRT-PCR) has been used to produce unique profiles of up-regulated and down-regulated gene expression in the liver of male Wistar rats following short term exposure to the non-genotoxic hepatocarcinogens, phenobarbital and WY-14,643. Animals were treated for 3 days, whereupon their livers were extracted and snap frozen. mRNA was prepared from the livers and used for ddRT-PCR. Individual bands from the differential displays were extracted and cloned. False positives were eliminated by dotblot screening and true positives then sequenced and identified.

INTRODUCTION

Safety evaluation of new chemicals usually necessitates the examination of genotoxic and carcinogenic potential using short-term in vitro and in vivo genotoxicity assays augmented by chronic bioassay tests. The short-term assays have proved useful in the early identification of potential genotoxic carcinogens, but their value is limited by observations which suggest that approximately 60% of chemicals identified as carcinogens in life-exposure studies produce mainly negative findings in short-term genotoxicity tests (1,2). Thus, there is currently no reliable and rapid means of evaluating the carcinogenic risk of new chemicals which fall into this latter group of compounds, termed non-genotoxic (or epigenetic) carcinogens.

Please send reprint requests to : Dr John Rockett, Molecular Toxicology Group, School of Biological Sciences, University of Surrey, Guildford, Surrey GU2 5XH, UK.

It is now evident that non-genotoxic carcinogens constitute a group of chemicals which are not only divergent in their interspecies toxicity, but also demonstrate different target organ selectivities and mechanisms of action (3,4). Elucidation of the molecular mechanisms underlying non-genotoxic carcinogenesis is currently underway, but the picture is still far from complete. It is anticipated that a better understanding of the early changes in genetic expression following exposure to non-genotoxic carcinogens will aid development of experimental strategies to identify cellular markers which are diagnostic for this type of toxicity.

Subtractive ddRT-PCR is a recently developed technique which facilitates the preferential amplification of gene products that demonstrate altered expression in target tissue(s) following exposure to chemical stimuli. Furthermore, using this technique, no prior knowledge of the specific genes which are up/down regulated is required. In the current study, we have undertaken to develop a specific and rapid assay for non-genotoxic carcinogens using the technique of ddRT-PCR. This has allowed us to identify characteristic

patterns of gene regulation following administration of two different non-genotoxic carcinogens (phenobarbital and Wy-14,643) and the subsequent identification of individual gene species which are regulated by this xenobiotic treatment.

MATERIALS AND METHODS

Animals and treatment

Phenobarbital (BDH, Poole, UK; 100 mg/kg/day) or [4-chloro-6-(2,3-xylidino)-2-pyrimidinylthio] acetic acid (Wy-14,643) (Campo, Emmerich; 250 mg/kg/day) was administered by gavage to groups of 3 male Wistar rats (150–200 g) on three consecutive days, whilst control animals received nothing. All animals had free access to food (rat and mouse standard diet, B&K Universal, Hull, UK) and water. The animals were killed on the fourth day, whereupon their livers were excised, sliced into 0.5 cm cubes, snap frozen in liquid nitrogen and then stored at -70°C.

mRNA extraction

Up to 0.25 g of each frozen liver sample was ground under liquid nitrogen using a mortar and pestle. mRNA was extracted from the ground liver using Promega's PolyATtract® System 1000 (Promega, Madison, WI, USA) according to the technical manual. The mRNA was DNase-treated (Promega, final concentration 10 U/ml) before phenol/chloroform extraction and ethanol precipitation. The mRNA was resuspended at a final concentration 500–1000 ng/µl.

ddRT-PCR

This was carried out using the PCR-Select™ cDNA Subtraction Kit (Clontech, Palo Alto, CA, USA) according to the manufacturer's instructions. Final PCR reactions were run on a 2% Metaphor agarose (FMC, Rockland, MD, USA) gel containing ethidium bromide (Sigma, Dorset, UK) and then overstained for 30 min with SYBR Green I DNA stain (FMC, 1:10 000 dilution in TAE).

Band extraction and cloning

Each discernible band from the differential display pattern was extracted from the gel with a scalpel and the DNA eluted using a Genelute™ Agarose Spin Column (Supelco, Bellefonte). An aliquot of the eluted DNA (5 µl) was re-amplified using the original ddRT-PCR nested primers and electrophoresed on a 2% agarose gel. The re-amplified band was extracted from the gel (as above) and the eluted DNA ligated directly into the TOPO TA Cloning® vector (Invitrogen, Carlsbad) before transformation in *Escherichia coli* TOP10F' One Shot™ cells (Invitrogen).

Stage 1 screening

Twelve transformed (white) colonies from each band were grown up for 6 h in 200 µl LB broth containing ampicillin (Sigma, 50 µg/ml) and 1 µl of this amplified by PCR reaction (as specified in ddRT-PCR technical manual). One quarter of the completed reaction was electrophoresed on a standard 2% agarose gel and one quarter on a 2% agarose gel containing HA Yellow (Hanse Analytik GmbH, Bremen, Germany, 1 U/µl) to discern the different cloning products. The remainder was used to prepare duplicate dotblots on Hybond N+ (nylon) membranes (Amersham, Little Chalfont, UK). Cultures containing different cloning products were grown up and a plasmid miniprep prepared from each (Wizard Plus SV Minipreps DNA Purification System, Promega) according to the manufacturer's instructions.

Stage II screening

The duplicate dotblots were probed with: (a) the final differential display reaction; and (b) the 'reverse-subtracted' differential display reaction. To make the 'reverse-subtracted' probe, the subtractive hybridisation step of the ddRT-PCR procedure was carried out using the original tester cDNA as a driver and the driver as a tester. Probing and visualisation were carried out using the ECL Direct Nucleic Acid Labelling and Detection System (Amersham) according to the manufacturer's instructions. Those clones which were positive for (a) but negative for (b), or showed a substantially larger positive signal with (a) compared to (b), were chosen for further analysis.

DNA sequencing

Positive clones as identified above were sequenced on an automated ABI DNA sequencer (Applied Biosystems, Warrington, UK).

A B

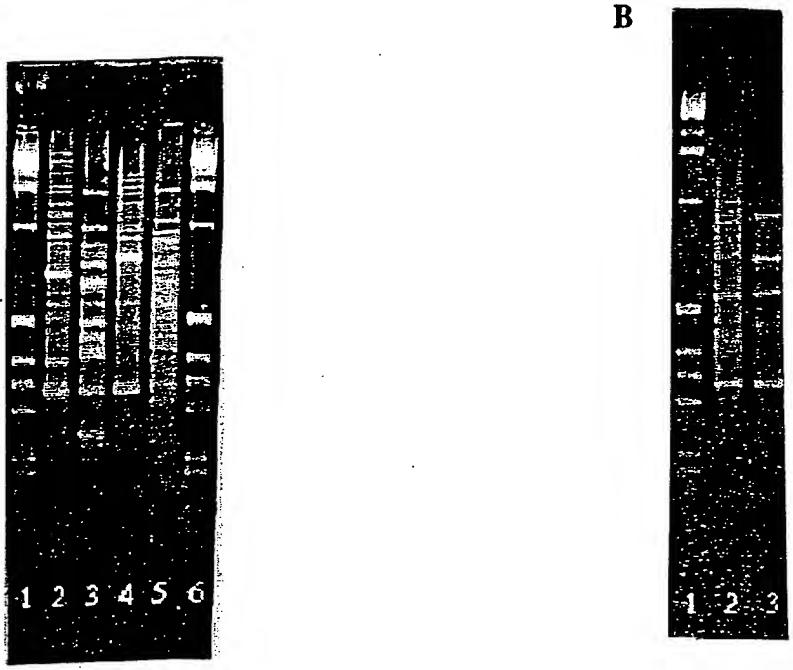


Fig. 1 : (*A*) Subtractive ddRT-PCR patterns obtained from rat liver following 3-day treatment with WY-14,643 or phenobarbital. Lane 1, 1 kb ladder; lane 2, genes up-regulated following Wy,14-643 treatment; lane 3, genes down-regulated following Wy,14-643 treatment; lane 4, genes up-regulated following phenobarbital treatment; lane 5, genes down-regulated following phenobarbital treatment; and lane 6, 1kb ladder. (**B**) Subtractive ddRT-PCR patterns obtained from rat liver showing relative changes when phenobarbital treated mRNA is subtracted from Wy-14,643-treated mRNA and vice-versa. Lane 1, 1 kb ladder; lane 2, genes showing increased expression following Wy-14,643 treatment compared to phenobarbital treatment; lane 3, genes showing increased expression following phenobarbital treatment compared to Wy-14,643 treatment. See Materials and Methods for further details.

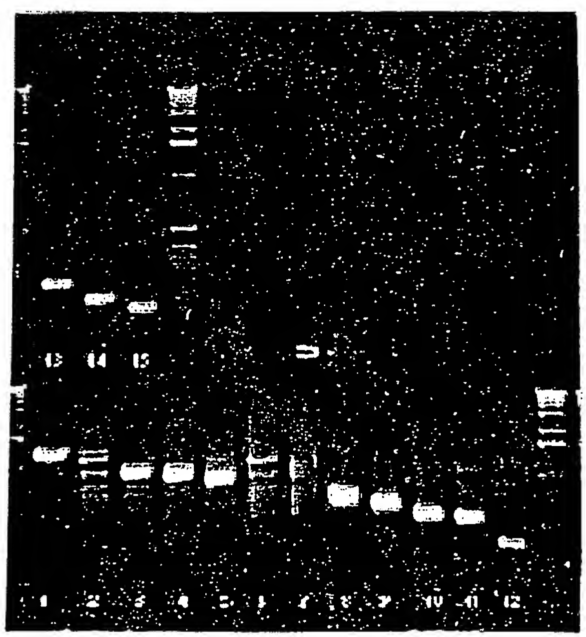


Fig. 2 : Re-amplified ddRT-PCR products which were down-regulated following phenobarbital treatment (upregulated bands were also re-amplified but gel not shown). Individual DNA bands excised from gel of ddRTR-PCR reactions were extracted, re-amplified and run on agarose gels to confirm amplification of correct band (numbered). See Materials and Methods for further details.

Table I : Rat liver genes down-regulated by phenobarbital treatment

Band number (Fig. 2) *(Approximate size in bp)*	*Phenobarbital down-regulated* *Highest sequence homology*		*FASTA-EMBL gene identification*
1 (1500)		95.3%	Rat mRNA for 3-oxoacyl-CoA thiolase
2 (1200)		92.3%	Rat hemopoxin mRNA
3 (1000)		91.7%	*R. rattus* alpha-2u-globulin mRNA
7 (700)	Clone 1	77.2%	*M. musculus* mRNA for Cl inhibitor
	Clone 2	94.5%	Rat electron transfer flavoprotein
	Clone 3	91.0%	Mouse topoisomerase 1 (Topo 1) mRNA
8 (650)	Clone 1	86.9%	Soares 2NbMT *M. musculus* (EST)
	Clone 2	96.2%	Rat alpha-2u-globulin (s-type) mRNA
9 (600)	Clone 1	86.9%	Soares mouse NML *M. musculus* (EST)
	Clone 2	82.0%	Soares p3NMF19.5 *M.musculus* (EST)
10 (550)		73.8%	Soares mouse NML *M. musculus* (EST)
11 (525)		95.7%	NCI_CGAP_Pr1 *H. sapiens* (EST)
12 (375)		100.0%	*R. norvegicus* mRNA for ribosomal protein
13 (230)	Clone 1	97.2%	Soares mouse embryo NbME135 (EST)
	Clone 2	100.0%	Rat fibrinogen B-beta-chain
	Clone 3	100.0%	Rat apolipoprotein E gene
14 (170)		96.0%	Soares p3NMF19.5 *M. musculus* (EST)
15 (140)		97.3%	Stratagene mouse testis (EST)
Others: (300)		96.7%	*R. norvegicus* RASP 1 mRNA
(275)		93.1%	Soares mouse mammary gland (EST)

EST = expressed sequence tag.
Bands 4–6 were shown to be false positives by dotblot analysis and, therefore, not sequenced.

Table II : Rat liver genes up-regulated by phenobarbital treatment

Band number *(Approximate size in bp)*	*Phenobarbital up-regulated* *Highest sequence homology*		*FASTA-EMBL gene identification*
5 (1300)		93.5%	Rat cytochrome P450IIB1
7 (1000)		95.1%	mRNA for rat preproalbumin
			Rat serum albumin mRNA
8 (950)		98.3%	NCI_CGAP_Pr1 *H. sapiens* (EST)
10 (850)		95.7%	Rat cytochrome P450IIB1
11 (800)	Clone 1	94.9%	Rat cytochrome P450IIB1
	Clone 2	75.3%	Rat cytochrome p450-L (p450IIB2)
12 (750)		93.8%	Rat TRPM-2 mRNA
			Rat mRNA for sulfated glycoprotein
15 (600)		92.9%	mRNA for rat preproalbumin
			Rat serum albumin mRNA
16 (550)	Clone 1	95.2%	Rat cytochrome P450IIB1
	Clone 2	93.6%	Rat haptoglobulin mRNA partial alpha
21 (350)		99.3%	*R. norvegicus* genes for 18S, 5.8S & 28S rRNA

EST = expressed sequence tag.
Bands 1–4, 6, 9, 13, 14 and 17–20 shown to be false positives by dotblot analysis and, therefore, not sequenced.

Identification of differentially-regulated genes

Gene-sequences were identified using the FASTA programme (http://www.ebi.ac.uk/htbin/fasta.py?request) to search all EMBL databases for matching DNA sequences.

RESULTS

Figure 1A,B shows the ddRT-PCR patterns of genes showing altered expression in rat liver following 3 day treatment with phenobarbital or Wy-14,643. Individual bands were isolated from the phenobarbital-modulated patterns (both up- and down-regulated), re-amplified (Fig. 2), cloned, screened for false positives and then identified. Those xenobiotic-modulated gene products identified to date are listed in Tables I and II.

DISCUSSION

The advent of combinatorial chemistry has led to the synthesis of millions of new chemical compounds, many of which may be potentially useful in pharmaceutical, agricultural or industrial applications. However, whilst there are tests available for those posing a genotoxic activity, there remains no short-term assay able to identify those chemicals which may belong to the non-genotoxic group of carcinogens.

We have used an adaptation of the subtractive hybridisation method – ddRT-PCR – to produce characteristic profiles or 'fingerprints' of those genes which are up-regulated or down-regulated in male rat liver following acute exposure to test chemicals. The ddRT-PCR profiles are characteristic and unique for each of the 2 compounds studied to date.

A number of those gene species showing altered expression following phenobarbital treatment have been cloned and identified (Tables I & II). It is interesting to note the presence of CYP2B2 in the up-regulated genes. This would, of course, be expected following exposure to phenobarbital and serves as a positive control for the method. Other genes which one might normally expect to be up-regulated do not appear in the table. However, it should be noted that not all bands seen on the differential display were extracted and re-amplified due to their being too faint or too close to other bands to accurately excise. Furthermore, it has been well documented [(5) and references therein] that a single band extracted from a differential display often represents a composite of heterogeneous products. We are currently examining new methods to: (i) improve resolution of the differential display patterns (including 2-D agarose gels); and (ii) distinguish those ddRT-PCR products which are identical in size, but different in sequence.

Our future efforts will be directed towards determining the extent of modulation of a number of the genes reported herein using semi-quantitative RT-PCR. This should reveal the extent of changes in expression of key gene products which may be involved in non-genotoxic hepatocarcinogenesis and thus help increase understanding of this process. Furthermore, it is anticipated that aligning ddRT-PCR profiles of different non-genotoxic agents found in responsive and non-responsive species may enable identification of those genes which are mechanistically relevant to the non-genotoxic hepatocarcinogenic process. Accordingly, this approach lends itself well to the identification, characterisation and sub-classification of possible different classes of non-genotoxic carcinogens.

ACKNOWLEDGEMENT

This work was funded by Rhône-Poulenc Agrochemicals, France

REFERENCES

1. Parodi S. (1992) : Non-genotoxic factors in the carcinogenic process: problems of detection and hazard evaluation. Toxicol. Lett., 64/65, 621-630.
2. Ashby J. (1992) : Prediction of non-genotoxic carcinogenesis. Toxicol. Lett., 64/65, 605-612.
3. Grasso G. and Sharratt M. (1991) : Role of persistent, non-genotoxic tissue damage in rodent cancer and relevance to humans. Annu. Rev. Pharmacol. Toxicol., 31, 253-287.
4. Lake B. (1995) : Mechanisms of hepatocarcinogenicity of peroxisome-proliferating drugs and chemicals. Annu. Rev. Pharmacol. Toxicol., 35, 483-507.
5. Smith N.R., Li A., Aldersley M., High A.S., Markham A.E., Robinson P.A. (1997) : Rapid determination of the complexity of cDNA bands extracted from DDRT-PCR polyacrylamide gels. Nucleic Acids Research 25 (17). 3552-3554.

JOURNAL OF

DRUG METABOLISM

A ... COKINETICS

PUBLISHED QUARTERLY

ELSEVIER

Toxicology 144 (2000) 13–29

www.elsevier.com/locate/toxicol

Use of suppression-PCR subtractive hybridisation to identify genes that demonstrate altered expression in male rat and guinea pig livers following exposure to Wy-14,643, a peroxisome proliferator and non-genotoxic hepatocarcinogen

John C. Rockett [1], Karen E. Swales, David J. Esdaile [2], G. Gordon Gibson *

Molecular Toxicology Group, School of Biological Sciences, University of Surrey, Guildford, Surrey GU2 5XH, UK

Abstract

Understanding the genetic profile of a cell at all stages of normal and carcinogenic development should provide an essential aid to developing new strategies for the prevention, early detection, diagnosis and treatment of cancers. We have attempted to identify some of the genes that may be involved in peroxisome-proliferator (PP)-induced non-genotoxic hepatocarcinogenesis using suppression PCR subtractive hybridisation (SSH). Wistar rats (male) were chosen as a representative susceptible species and Duncan–Hartley guinea pigs (male) as a resistant species to the hepatocarcinogenic effects of the PP, [4-chloro-6-(2,3-xylidino)-2-pyrimidinylthio] acetic acid (Wy-14,643). In each case, groups of four test animals were administered a single dose of Wy-14,643 (250 mg/kg per day in corn oil) by gastric intubation for 3 consecutive days. The control animals received corn oil only. On the fourth day the animals were killed and liver mRNA extracted. SSH was carried out using mRNA extracted from the rat and guinea pig livers, and used to isolate genes that were up and downregulated following Wy-14,643 treatment. These genes included some predictable (and hence positive control) species such as CYP4A1 and CYP2C11 (upregulated and downregulated in rat liver, respectively). Several genes that may be implicated in hepatocarcinogenesis have also been identified, as have some unidentified species. This work thus provides a starting point for developing a molecular profile of the early effects of a non-genotoxic carcinogen in sensitive and resistant species that could ultimately lead to a short-term assay for this type of toxicity.

Keywords: Wy-14,643; Peroxisome proliferator; Non-genotoxic hepatocarcinogenesis; Suppression PCR subtractive hybridisation; RT-PCR; Rat; Guinea pig; Gene regulation; Differential gene display; Gene profiling

* Corresponding author. Tel.: +44-1483-259704; fax: +44-1483-576978.
E-mail address: g.gibson@surrey.ac.uk (G.G. Gibson)
[1] Present address: US Environmental Protection Agency, National Health and Environmental Effects Research Laboratory, Reproductive Toxicology Section, Research Triangle Park, NC 27711, USA.
[2] Present address: Rhone-Poulenc Agrochemicals, Toxicology Department, Sophia-Antipolis, Nice, France.

0300-483X/00/$ - see front matter
PII: S0300-483X(99)00214-0

Introduction

Γhe advent of combinatorial chemistry and nputer-aided drug design has led to a recent surge in the number of chemical compounds ıt have potential therapeutic, agricultural and lustrial applications. Although it has been sugted that the contribution of synthetic chemicals the overall incidence of human cancer is low, re still remains an absolute requirement to aluate all new chemicals for toxic and carcinoıic potential. The latter is one of the most ›blematic areas of chemical safety evaluation l is usually carried out using short-term in vitro l in vivo genotoxicity assays augmented by onic bioassay tests. The short-term assays have ›ved useful in the early identification of poten genotoxic carcinogens, but their value is liml by observations that suggest that ›roximately 60% of chemicals identified as carɔgens in life-exposure studies produce mainly ative findings in short-term genotoxcity tests hby, 1992; Parodi, 1992). Thus, there is curtly no reliable and rapid means of evaluating carcinogenic risk of new chemicals that fall › this latter group of compounds, termed nonotoxic (or epigenetic) carcinogens.

)ne approach to addressing this problem is to :idate the molecular mechanisms by which ›wn non-genotoxic carcinogens act. It should ı be possible to identify common factors/ :hanisms that can serve as early biomarkers of :inogenic potential for new chemicals. To this , a large number of groups have reported on various effects of non-genotoxic compounds various animal species (Marsman et al., 1988; :e et al., 1993; Cattley et al., 1994; Hayashi et 1994; Human and Experimental Toxicology, 4; Anderson et al., 1996). However, the mechtic picture is still far from complete with many hose genes involved in the carcinogenic proremaining unknown, and their identification efore remains a key goal in elucidating the ecular mechanisms by which non-genotoxic inogenesis occurs.

ıbtractive hybridisation (SH) and related tech›gies such as representational difference analy(RDA) (Hubank and Schatz, 1994) and differential display (DD) (Liang and Pardee, 1992) can be used to aid the isolation of genes showing altered expression in target tissues following exposure to a chemical stimulus. These techniques can also be used to identify differential gene expression in neoplastic and normal cells (Liang et al., 1992), infected and normal cells (Duguid and Dinauer, 1990), differentiated and undifferentiated cells (Sargent and Dawid, 1983; Guimaraes et al., 1995), activated and dormant cells (Gurskaya et al., 1996; Wan et al., 1996), different cell types (Hedrick et al., 1984; Davis et al., 1984) amongst others. Most importantly, using such approaches, no prior knowledge of the specific genes that are upregulated/downregulated is required.

Using a variation of SH, termed suppression-PCR subtractive hybridisation (SSH) (Diatchenko et al., 1996), we have previously reported the isolation of a number of genes showing altered expression in male rat liver following acute exposure to phenobarbital (Rockett et al., 1997). In the current work we have used the same experimental approach to isolate genes that are differentially expressed in the livers of male rats and guinea pigs following short-term (3-day) exposure to the peroxisome proliferator (PP) and non-genotoxic hepatocarcinogen, Wy-14,643. We have isolated and identified a number of gene species, some of which may be important in the induction of, or protection against, non-genotoxic hepatocarcinogenesis.

2. Materials and methods

2.1. Animals and treatment

All animal experiments were undertaken in accordance with Her Majesty's Home Office Department guidelines under the auspices of approved personal and project licences. Male Wistar rats (150–200 g) and male Duncan–Hartley guinea pigs (250–300 g) were obtained from Kingman and Bantam (Hull, UK). Upon receipt, both groups were randomly assigned into two groups of four. They were maintained on a rat, mouse or guinea pig standard diet (B&K Univer-

sal, Hull) and a daily cycle of alternating 12-h periods of dark and light. The room temperature was maintained at 19°C and a relative humidity of 55%. The animals were acclimatised to this environment for 7 days before treatment commenced. [4-chloro-6-(2,3-xylidino)-2-pyrimidinylthio] acetic acid (Wy-14,643, Campo, Emmerich; 250 mg/kg per day in corn oil) was administered by gavage to the treated groups of rats and guinea pigs on 3 consecutive days, whilst control groups received an equal volume of corn oil only. During this time, all animals had free access to food and water. The animals were killed by cervical dislocation on the fourth day, and their livers immediately excised, weighed, sliced into approximately 0.5-cm cubes, snap frozen in liquid nitrogen and stored at -70°C.

2.2. mRNA extraction

Approximately 0.25 g of each frozen liver sample was ground under liquid nitrogen using a mortar and pestle. Messenger RNA was extracted from the ground liver using the PolyATtract® System 1000 kit (Promega, Madison, USA) according to the technical manual provided by the manufacturers. The mRNA was DNase-treated (RQ Rnase-free Dnase, Promega, final concentration 10 U/ml) before phenol/chloroform extraction and ethanol precipitation. The mRNA was redissolved at a final concentration 500–1000 ng/μl.

2.3. cDNA Subtraction

This was carried out using the PCR-Select™ cDNA Subtraction Kit (Clontech, Palo Alto, USA) according to the manufacturer's instructions. Subtractions were carried out with mRNAs derived from single animals. The mRNA from the remaining three animals in each group was later used for quantitative RT-PCR analysis of specific genes.

2.4. Band extraction and cloning

The secondary PCR reactions from the cDNA subtraction procedure were run on a 2% Metaphor agarose gel (FMC, Rockland, USA) containing 0.5 μg/ml ethidium bromide (Sigma, Dorset, UK). One times TAE (0.04 M Tris-acetate, 0.001 M EDTA) was used to prepare the gel and as the running buffer. After running for 6–7 h at 3.75 V/cm, the gel was overstained for 30 min with SYBR Green I DNA stain (FMC, 1:10 000 dilution in 1 × TAE). Each discernible band of the differential display pattern was extracted from the gel with a scalpel and the DNA eluted using a Genelute™ agarose spin column (Supelco, Bellefonte, USA). Five microlitres of the eluted DNA was reamplified using the original nested (secondary) PCR primers supplied with the PCR-Select™ cDNA subtraction kit. The PCR products were electrophoresed on a 2% standard agarose gel (Boehringer Mannheim, East Sussex, UK) and the reamplified target bands extracted from the gel as above. The eluted DNA was immediately ligated into a TOPO TA Cloning® vector (Invitrogen, Carlsbad, USA) before transformation in *Escherichia coli* TOP10F' One Shot™ cells (Invitrogen).

2.5. Colony screening

2.5.1. Stage I

Eight transformed (white) colonies from each band were grown up for 6 h in 200 μl LB broth containing ampicillin (Sigma, 50 mg/ml). One microlitre of this was subjected to PCR using the same conditions and nested primers as described above. One tenth (2 μl) of the completed PCR reaction was electrophoresed on a 2% standard agarose gel and one tenth on a 2% standard agarose gel containing HA red (Hanse Analytik GmbH, Bremen, Germany, 1 U/ml) to discern the differentially cloned products. The remainder of the PCR reaction was used to prepare duplicate dotblots on Hybond N^+ membranes (Amersham, Little Chalfont, UK).

2.5.2. Stage II

The duplicate dotblots were probed with (a) the final differential display reaction and (b) the 'reverse-subtracted' differential display reaction. To make the 'reverse-subtracted' probe, the subtractive hybridisation step of the differential display

.T-PCR procedure was carried out using the riginal tester (treated) mRNA as the driver and ie original driver (control) mRNA as the tester. robing and visualisation were carried out using ie ECL direct nucleic acid labelling and detecon system (Amersham, Little Chalfont, UK) acording to the manufacturer's instructions. Those .ones that were positive for (a) but negative for), or showed a substantially larger positive sigal with (a) compared to (b), were selected for NA sequence analysis.

.6. *DNA sequencing*

The remainder of the cultures (prepared in age 1 screening) containing different cloning roducts (as discerned in the two screening steps) ere grown up overnight in 5 ml LB broth conining ampicillin (50 mg/ml). A plasmid miniprep as prepared from each (Wizard Plus SV linipreps DNA purification system, Promega) cording to the manufacturer's instructions. The oned inserts were sequenced on an automated BI DNA sequencer (Applied Biosystems, Warngton, UK) using the M13 forward primer 3TAAAACGACGGCCAGT) or M13 reverse imer (AACAGCTATGACCATG).

7. *Identification of differentially regulated genes*

Gene sequences thus obtained were identified sing the FASTA 3.0 programme (Lipman and earson, 1985; Pearson and Lipman, 1988) (http:/ /ww.ddbj.nig.ac.jp/E-mail/homology.html) to arch all EMBL databases for matching DNA quences. Each clone sequence was submitted in e forward and reverse direction, and the one turning the highest statistical probability of atch to a known sequence was noted. Sequence omologies between our submitted clone sequence id the queried database sequence were determined (by FASTA) over a region of at least 60 ise pairs.

8. *RT-PCR analysis of selected candidate genes*

cDNA sequences of the target genes were obined from the NIH gene database (GenBank at http://www.ncbi.nlm.nih.gov/Web/Search/index.html) and the computer programme GENE JOCKEY (BioSoft, Cambridge, UK) used to select primer pairs from these sequences. Where guinea pig sequences were available, rat and guinea pig sequences were aligned and primers chosen from regions of homology. If guinea pig sequences were not available, rat and human sequences were used. In cases where exact homology could not be found, the sequence from the rat was used. In the case of CD81 only, no rat or guinea pig sequences were available and so mouse and human sequences were aligned and a primer pair chosen from a region of homology. Primers (obtained from Gibco-BRL, Paisley, UK) were dissolved at a concentration of 50 pmol/μl in sterile distilled water and stored at −20°C. The primer pairs used plus other reaction parameters are shown in Table 1. mRNA was extracted (as described above) from all four treated animals and from three animals in the control group. Integrity of the eluted mRNA was confirmed on a 2% agarose gel, and the concentration and purity were measured using a Genequant II spectrophotometer (LKB, Bromma, Sweden) and then diluted to 10 ng/μl. One microlitre of this latter solution was used per RT-PCR reaction.

RT-PCR was carried out in a single tube (50 μl) reaction using the Access RT-PCR system (Promega) according to manufacturer's instructions. In the kinetic and quantitative analyses, omission of RNA was used as a control for the presence of any contaminating DNA. After obtaining a PCR signal of the correct size and optimising the reaction conditions, each PCR product was digested with between two and four separate restriction enzymes. Specific restriction patterns were thus obtained, which further confirmed the identity of the PCR products as being the original target genes. Kinetic analysis (14–32 cycles) was then performed in each case to determine the location of the mid-log phase.

For the semi-quantitative analysis of each target gene, RT-PCR reactions were carried out in triplicate for each sample to reduce the effect of intertube RT-reaction variations (Kolls et al., 1993) and pipetting errors. For each gene, a mastermix containing enough reagents for three times

Table 1
Primer sequences and reaction conditions used in semi-quantitative RT-PCR analysis of selected genes

Transcript	Genbank accession No.	Primer sequences		Size of rat PCR product (bp)	Annealing temperature (°C) rat/guinea pig	No. of PCR cycles rat/guinea pig
		upstream	downstream			
Albumin	J00698 (rat)	TGGAGAGA-GAGC-CTTCAAAGC	CTTAG-CAAGTCTCAGCAGCA	436	60/59	15/22
Bifunctional enzyme	K03249 (rat)	GCACC-CACTTCTTCT-CACCAGC	TGGCAATGATG-GTCCAGTAAGG	347	57/–	21/–
CYP2C11	J02657 (rat)	CCATCATGACC-CTGAGG	GAAGTCCCGAG-GATTGT	410	50/–	20/–
CYP4A1	M14972 (rat)	GATGGCTGCAC-CATGAG	GGCCTTTG-GATCTGATC	357	57/–	22/–
Catalase	M11670 (rat)	ACCAAATACTC-CAAGGCAAAGG	GCCCTG-GTCAGTCTTG-TAATGG	450	63/–	27/–
CD81 (TAPA-1)	X59047 (mouse)	ATTTCGTCTTCTGGCTGGCTGG	GCCTGGTCATA-GAACTGCTTCA	337	57/59	23/22
Contrapsin-like protease inhibitor	RNCCP23 (rat)	GACTATGTGAG-CAATCAGAC	GTCTCTGGTTG-CAAGCT	341	50/–	20/–
Parathymosin-α (Zn^{2+} binding protein)	X64053 (rat)	CGGCACCAT-GTCGGAGAAGA	TTGTGTGTTCCT-GCCCCACC	382	62/–	24/–
Transferrin	D38380 (rat)	AGCTGTGT-CAACTGT-GTCCAGG	GAGGAGAGCC-GAACAGTTG-GAA	360	57/59	22/22
UDP-GT	U06273 (rat)	GGAT-GTCTGGGAAGTGG	GCAGTTCAGC-TATCAGCT	495	50/–	23/–
DownUnknown-1	n/a	CGACGTTTC-CAAGGCA	TGTTGCGGCA-GAGTGGA	318	55/–	25/–
Znα2glycoprotein	D21058 (rat)	CAAATAACA-GAAGCAGTG-GAGC	GACTTCCAC-CTCCATCCAGG	433	57/–	23/–

the number of samples (seven for rat, six for guinea pig) was prepared except that mRNA was omitted, the latter being added after aliquoting 49 μl of the mastermix into an appropriate number of tubes. Amplification of albumin (the reference gene) was carried out in separate tubes since the mid-log phase of this gene is at a much lower cycle number than the target genes due to its high abundance. All RT-PCR products were analysed on 2% agarose gels containing 0.5 μg/ml ethidium bromide. The target gene samples were loaded on the gel first and run in at 3 V/cm for 10 min. The corresponding albumin samples were then loaded and the gel run for a further 1/2 h. In this way, all RT-PCR products from each target gene and albumin from the corresponding samples could be run on the same gel. Gels were photographed using type 665 posi-neg film (Sigma) and quantitation of the band intensity was carried out using a dual wavelength flying spot laser scanner densitometer (Shimadzu).

2.9. Statistical analysis

Statistical analysis of unpaired samples was carried out using the two-tailed Student's t-test. Values were considered statistically significant at $P < 0.05$ or less.

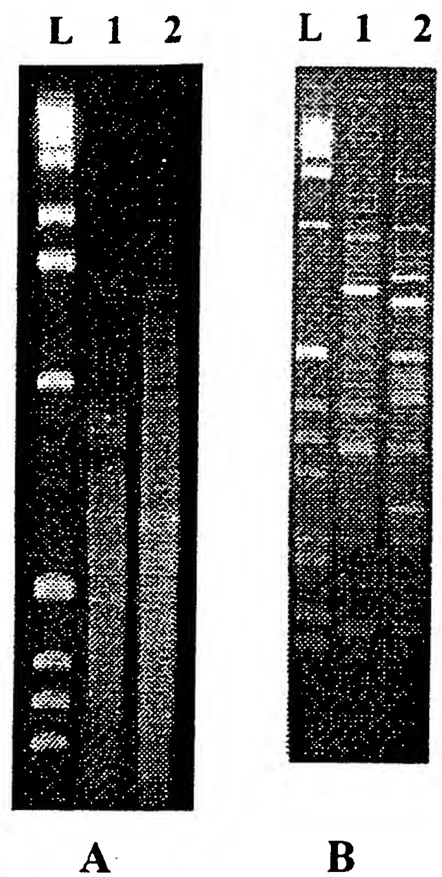


Fig. 1. Final displays of differentially expressed genes that were (1) upregulated and (2) downregulated in rat (A) and guinea pig (B) livers following 3-day treatment with Wy-14,643. mRNA extracted from control and treated livers was used to generate the differential displays using the PCR-Select cDNA subtraction kit (Clontech). Lane (L) is a 1 Kb DNA Ladder standard and 10 μl of secondary PCR reaction were loaded in all other lanes.

3. Results

3.1. Cloning and screening of transcripts

For both the rat and guinea pig experimental groups, cDNA subtraction was carried out in the forward (control driving tester) and reverse (tester driving control) directions to isolate both upregulated and downregulated mRNA species respectively. Using a standard primary hybridisation time of 8 h we obtained a substantial amount of non-specific products in all the final differential displays (data not shown). This background smearing was almost completely removed by reducing the primary hybridisation time to 4 h (CLONTECHniques, 1996). Fig. 1 shows the ddRT-PCR patterns of genes showing altered expression in rat and guinea pig liver following 3-day treatment with Wy-14,643. The profiles are unique for each species, and in each case the profile for the upregulated genes (control mRNA driving tester mRNA) is different to that obtained for the downregulated genes (tester mRNA driving control mRNA).

The practical outcome of the SSH method is that a series of differentially expressed genes is observed as a ladder on an agarose gel. The majority of these gene fragments fall within the 150–2000 bp range, with bands up to 5 Kbp occasionally being observed. Each band may theoretically consist of one or more products of similar size, as the gel has a maximum resolution

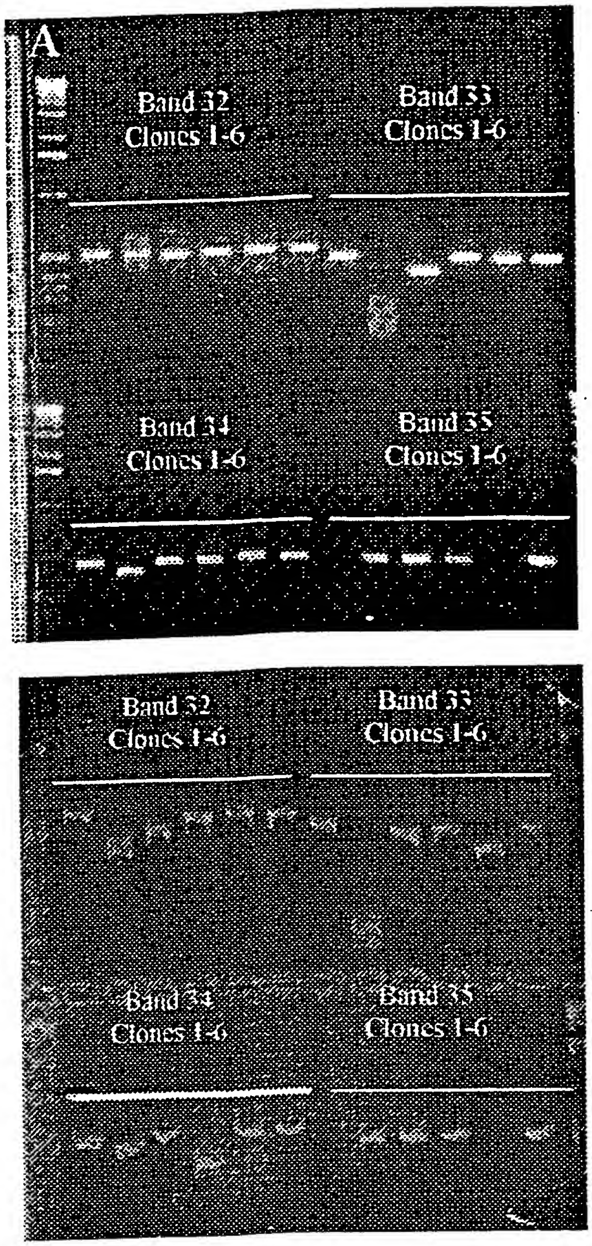


Fig. 2. Discrimination of different ddRT-PCR products having the same molecular size using HA-red. Gel (A) is a 2% standard agarose gel. Gel (B) is a 2% standard agarose gel containing 1 U/ml HA-red. Band numbers refer to the sequential bands (largest to smallest) extracted from the original display of genes upregulated in rat liver following 3-day treatment with Wy-14,643. Ten micorlitres of each PCR reaction were loaded per lane.

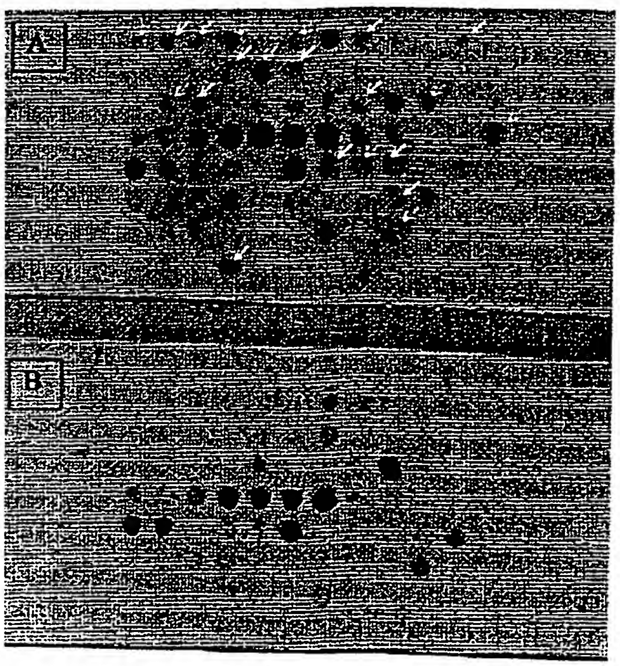
of approximately 1.5% (3 bp per 200). In addition, there may be two or more products that are the same size, but have a different sequence. Therefore some form of discrimination must be employed to isolate as many of these products as possible. HA-red screening (Geisinger et al., 1997) of a number of clones derived from each band provided a means to discriminate between different gene species of the same size. A typical example of such a gel is shown in Fig. 2. In total, 88 and 48 apparently different clones were obtained from the final differential expression patterns of upregulated and downregulated rat genes, respectively. Sixty nine and 89 apparently different clones were obtained from the final differential expression patterns of the upregulated and downregulated guinea pig genes, respectively.

Having identified as many different candidate gene products as possible in the screening step I, a second screening step was carried out on every clone to confirm those that represented true differentially expressed genes. This is necessary since no subtraction technique is 100% efficient. The approach we used, termed PCR-select differential screening (as recommended in Clontech's PCR-select cDNA subtraction kit protocol), utilises the forward and reverse subtractions as an aid to screening for the true differentially expressed genes (CLONTECHniques, 1997). Because these probes have already undergone subtraction, they have been enriched for differentially expressed genes and are therefore more sensitive than unsubtracted driver/tester cDNA probes for detecting true differential expression. All the clones that were isolated from each display were dotblotted and probed with the display from which they was obtained, plus the corresponding reverse-subtracted display. An example of such a blot is shown in Fig. 3. Clones corresponding to authentic differentially expressed mRNAs hybridised with the subtracted cDNA probe, but not the reverse-subtracted probe. We also included in the authentic positives, those clones that gave a substantially greater signal with the subtracted probe compared to the reverse-subtracted probe. False positives hybridised with either both probes or with neither probe. Of the original 88 upregulated and 48 downregulated rat clones selected for this screening step, 28 (32%) and 15 (31%) respectively, were found to be true positives. In the rat,

3 (100%) of the true positive upregulated genes (Table 2) and 11 (73%) of the true positive downregulated genes (Table 3) were non-redundant. Of the original 69 upregulated and 89 downregulated guinea pig clones selected for this screening step, 48 (70%) and 37 (42%) respectively, were found to be true positives. Thirty six (75%) of the upregulated genes (Table 4) and 33 (89%) of the downregulated genes (Table 5) were non-redundant.

2. Identification of clones

On sequence analysis it was found that some clones were unsequencable in the first instance (M13 forward primer) due to long polyA runs that appeared to prematurely terminate the sequencing reaction. These clones were therefore sequenced from the opposite direction using the M13 reverse primer. Those xenobiotic-modulated gene products identified to date are listed in Tables 2 and 3 (rat) and Tables 4 and 5 (guinea pig).



. 3. Dot blots of clones of putative upregulated gene species ated from guinea pig liver following 3-day treatment with -14,643. All clones identified in the stage I screening step methods) were blotted and probed with (A) the differen- display from which they originated (control driving .ted) and (B) the reverse subtraction (treated driving con-). Arrows indicate some of the true differentially expressed ıes.

Table 2
Identification of genes that were upregulated in male rat liver following 3-day treatment with WY-14,643

FASTA-EMBL gene identification (rat unless otherwise stated)	Accession No.	Sequence homology[a] (%)
Carnitine octanoyl transferase	RN26033	99
NCI_CGAP_Li1 (*H. sapiens*) (EST[b])	HS1275949	98
Peroxisomal enoyl hydratase-like protein	RN08976	98
Liver fatty acid binding protein	V01235	96
Soares mouse p3NMF19.5 *M. musculus* cDNA clone	AA038051	96
Cytochrome p450IVA1	RNCYPLA	94
Mit. 3-hydroxyl-3-methylglutaryl CoA synthase	RNHMGCOA	94
Rabgeranylgeranyl transferase component B	RNRABGERA	94
Genes for 18S, 5.8S, and 28S ribosomal RNAs	RNRRNA	94
Carnitine acetyl transferase (mouse)	MMRNACAR	92
Soares mouse NML (EST)	MM1157113	92
Bone marrow stromal fibroblast (*H. sapiens*) cDNA clone HBMSF2E4 (EST)	AA545726	92
7.5dpc embryo (mouse) (EST)	AA408192	92
Alpha-1-macroglobulin	RNALPH1M	91
Transferrin	RNTRANSA	91
Lecithin:cholesterol acyltransferase	RNU62803	90
Zn-α2-glycoprotein	RNZA2GA	90
Serum albumin	RNJALBM	89
Fructose-1,6-bisphosphate 1-phosphohydrolase	RNFBP	88
Soares mouse melanoma (EST) (S[c])	AA124706	88
Soares mouse 3NbMS (EST) (AS[c])	AA154039	88

Table 2 (*Continued*)

FASTA-EMBL gene identification (rat unless otherwise stated)	Accession No.	Sequence homology[a] (%)
17-β-hydroxsteroid dehydrogenase	RN17BHDT2	87
Soares mouse p3NMF19.5 (EST)	AA038051	87
Peroxisomal enoyl-CoA:hydratase -3-hydroxyacyl CoA bifunctional enzyme	RNPECOA	85
Integral membrane protein, TAPA-1 (CD81) (mouse)	S45012	81
Soares mouse lymph node (EST)	MMAA88445	81
H. sapiens (clone zap128) mRNA	L40401	76
Lysophospholipase homologue (human)	HSU67963	76
Soares mouse lymph node (EST)	AA217044	74

[a] Refers to the nucleotide sequence homology between the cloned band isolated from the differential display and the corresponding gene derived from the EMBL gene sequence bank.
[b] EST is 'expressed sequence tag' — a gene of as yet unknown identity and function.
[c] Where sequence homologies were equal in both directions of the isolated band, both the sense (S) and antisense (A) identities are given.

Table 3
Identification of genes that were downregulated in male rat liver following 3-day treatment with Wy-14,643

FAST-EMBL gene identification (rat unless otherwise stated)	Accession No.	Sequence homology[a] (%)
NCI_CGAP_Li1 (*H. sapiens*) (EST[b])(S[c])	AA484528	99
NCI_CGAP_Pr1 (*H. sapiens*) (EST)(AS[c])	AA469320	99
UDP-glucuronosyl-transferase (UGT2B12)	RN06273	98
Complement component c3	RNC3	96
Soares mouse placenta (S)	AA023305	96
Ape (chimpanzee) 28S rRNA (AS)	PTRGMC	96
Rat CYP2C11	RNCYPM1	95
Ribosomal protein S5	RNRPS5	94
Transthyretin	RNTTHY	94
Contrapsin-like protease inhibitor	RNCCP23	89
Prostaglandin F2a (S)	RN26663	84
β-2-microglobulin (AS)	RNB2MR	84
Apolipoprotein C-III	RNAPOA02	82
Parathymosin-alpha (zinc2+-binding protein)	RN11ZNBP	75

[a] Refers to the nucleotide sequence homology between the cloned band isolated from the differential display and the corresponding gene derived from the EMBL gene sequence bank.
[b] EST is 'expressed sequence tag' — a gene of as yet unknown identity and function.
[c] Where sequence homologies were equal in both directions, both the sense (S) and antisense (A) identities are given.

In all cases, both the forward and reverse sequence of the target clones were analysed and the gene having the highest statistical homology noted.

3.3. RT-PCR analysis of selected clones

The results of a typical RT-PCR semi-quantitation experiment for transferrin in the rat is given in Fig. 4 and the results for a total of 12 selected genes in both the rat and guinea pig are shown in Table 6.

4. Discussion

It is now apparent that all cancers arise from accumulated genetic changes within the cell. Although documenting and explaining these changes presents a formidable obstacle to understanding the different mechanisms of carcinogenesis, the experimental methodology is now available to begin attempting this difficult challenge. In order to begin the elucidation of the molecular mechanisms involved in non-genotoxic hepatocarcino-

ıesis, we have used SSH to identify a number of ıes that are upregulated or downregulated in ıle rat and guinea pig livers following short m exposure to the PP, Wy-14,643. We have ɜd the rat model to represent a species suscepti- : to the non-genotoxic carcinogenic effect of 's and the guinea pig as a resistant species rton et al., 1984; Rodricks and Turnbull, 1987; Lake et al., 1989; Makowska et al., 1992; Lake et al., 1993).

Gurskaya et al. (1996), who originally developed the SSH technique, cloned the products of the secondary PCR reaction and screened a small number of randomly selected colonies for differentially expressed clones using northern hybridisation. However, we decided against this approach

ɔle 4
ntification of genes that were upregulated in male guinea pig liver following 3-day treatment with WY-14,643

STA-EMBL gene identification (guinea pig unless otherwise stated)	Accession No.	Sequence homology[a] (%)
·boxylesterase	AB010634	97
nplement C3 protein (GPC3)	M34054	97
osolic aldehyde dehydrogenase (sheep)	U12761	92
alase (human)	X04076	89
ochondrial aspartate aminotransferase (pig)	M11732	89
ngation factor-1-alpha (rabbit)	X62245	88
I_CGAP_Br2 *H. sapiens* cDNA clone (EST) (Similar to chick mit. phosphoenolpyruate carboxykinase)	AA587436	87
ha-1-antiproteinase S	M57270	83
'ormyltetrahydrofolate dehydrogenase (rat)	M59861	83
osomal protein L6 (rat)	X87107	83
res pregnant uterus Nb (EST) (mouse)	AA156847	83
ochondrial citrate transport protein (human)	L77567	80
oplasmic chaperonin hTRiC5 (human)	U17104	80
·ha-1-antiproteinase F	M57271	77
erogeneous nuclear ribonuclearprotein c1/c2 (human)	D28382	77
res parathyroid tumour (EST) (similar to human serum albumin precursor)	AA860651	76
ıtagene mouse kidney (EST)	AA107327	75
res parathyroid tumour NbHPA human cDNA (EST)	AA860653	74
res mouse mammary gland (EST)	AA619297	74
NA clone 15 004 (EST) (human)	H01826	74
res senescent fibroblasts (EST) (mouse)	W52190	74
proalbumin (human)	E04315	72
NA clone 73 169 (EST) (human)	T56624	72
ımin D-binding protein (human)	L10641	71
›H gene (exon 8) (human)	Y11498	71
₹L flow sorted chromosome	B05457	71
res foetal liver spleen (EST) (mouse)	AA009524	71
res foetal heart NbMH19W (EST) (mouse)	AA009421	69
res foetal heart NbHH19W *H. sapiens* cDNA clone (EST)	W94377	67
nylalanine hydroxylase (human)	U49897	67
line-5-carboxylate dehydrogenase (human)	U24266	66
tathione-*S*-transferase homologue (human)	U90313	65
I_CGAP_GCBI (EST) (human)	AA769294	65
ıective protein (human)	M22960	64
ıe 27 375 (EST) (human)	N37046	62
tagene colon (#937 204) *H. sapiens* cDNA clone (EST)	AA149777	62

Refers to the nucleotide sequence homology between the cloned band isolated from the differential display and the correspond- gene derived from the EMBL gene sequence bank.

Table 5
Identification of genes that were downregulated in male guinea pig liver following 3-day treatment with WY-14,643

FASTA-EMBL gene identification (guinea pig unless otherwise stated)	Accession No.	Sequence homology[a] (%)
Complement C3 protein	M34054	97
Murinoglobulin	D84339	95
Alpha-1-antiproteinase F	M57271	88
Elongation factor-alpha-1 (rabbit)	X62245	89
Coupling protein G (human)	X04409	88
NCI_CGAP_Ov1 (EST[b]) (human)	AA586309	87
Lecithin:cholesterol acetyl transferase (rabbit)	D13668	85
Aldolase B (human)	X00270	84
Anti-thrombin III (human)	E00116	80
Phenylalanine hydroxylase (human)	K03020	80
Inter-α-trypsin inhibitor (human)	D38595	79
Normalised rat muscle (EST) (S[c])	AA849753	78
Normalised rat ovary (EST) (AS[c])	AA801059	78
Complement factor Ba fragment (human)	X00284	77
Dihydrodiol dehydrogenase (human)	U05598	76
Spot14 gene (thyroid-inducible hepatic protein)(human)	Y08409	75
BAC clone 174p12 (human)	AC004236	75
Mitochondrial aldehyde dehydrogenase (human)	X05409	74
Preproalbumin (human)	E04315	74
NCI_CGAP_Pr9 (EST) (human) (S)	AA533142	74
Normalised rat placenta (EST) (AS)	AA851197	74
Heparin sulfate proteoglycan (human)	J04621	73
cDNA clone 33 992 (EST) (human)	R24330	73
Retinol dehydrogenase (rat)	U33501	71
TAPA-1 integral membrane protein (CD81) (mouse)	S45012	71
Complement component c5s	M35525	70
Apolipoprotein B (pig)	L11235	69
cDNA clone 143 918 (EST) (human)	R76742	68
α-fibrinogen (human)	K02569	68
Soares foetal liver spleen 1NF (mouse)	W03726	68
Barstead bowel (EST) (mouse)	AA232049	67
UDP glucuronosyl transferase (cat)	AF0309137	66
Myeloid leukaemia cell differentiation protein (MCL-1) (human) (S)	L08246	65
STS SHGC-34 987 (human) (AS)	G27984	65
Soares mouse 3NME125	AA222798	64
Stratagene mouse embryonic (EST) (S)	AA199420	64
Rad 52 (mouse)	AF004854	63

[a] Refers to the nucleotide sequence homology between the cloned band isolated from the differential display and the corresponding gene derived from the EMBL gene sequence bank.
[b] EST is 'expressed sequence tag' — a gene of as yet unknown identity and function
[c] Where sequence homologies were equal in both directions, boththe sense (S) and antisense (A) identities are given.

for several reasons: (1) the kinetics of ligation and transformation favour the isolation of smaller PCR products, thereby producing a misrepresentation of larger gene products; (2) northern blot analysis is notoriously insensitive and is unlikely to confirm expression of rare transcripts; (3) there is no measurable end point to the screening of clones produced in this way other than to analyse every transformed colony. We used instead an alternative approach; after running out the differ-

ntial display on a high-resolution agarose gel Fig. 1) and overstaining with SYBR Green I to nhance visualisation, the composite bands were ndividually extracted, reamplified and cloned. However, it has been well documented that single bands from differential displays often contain a heterogeneous mixture of different products Mathieu-Daude et al., 1996; Smith et al., 1997). This is because polyacrylamide gels cannot discriminate between DNA sequences that differ in ize by less than about 0.2% (Sambrook et al., 989). High-resolution agarose gels such as those ised in this work are even less sensitive, normally only discriminating products that differ in size by t least 1.5%. The use of the HA-red screening tep enables resolution of identical or nearly identical sequences based on their AT content (Wawer t al., 1995) and is sensitive down to < 1% difference. Furthermore, it is rapid, technically simple nd does not require the use of radiolabels. Geisinger et al. (1997) originally demonstrated the usefulness of using HA-red to identify different products cloned from the same band of an RNA differential display experiment by simultaneously unning them in normal agarose (to discriminate by size) and in normal agarose containing HA-red to discriminate by AT content). We have found hat this approach is equally useful for identifying different gene species cloned from the same band of our SSH display.

Diatchenko et al. (1996) reported that SSH is highly efficient at producing differentially expressed gene species. However, we also included a econd screening step to further confirm that the lones isolated from the differential display were ndeed differentially expressed. Duplicate dotblots of the candidate clones were blotted with the display from which they were originally isolated nd with the 'reverse subtraction' display. To nake the reverse-subtracted probe, the subtractive hybridisation step of the procedure was carried out using the original tester cDNA as a driver, nd the original driver cDNA as a tester. In this way, clones that are false positives can be identified through their presence in both blots. Such alse positives most commonly arise through having a very high abundance in the initial sample or nusual hybridisation properties (Li et al., 1994).

Although the SSH method itself has been shown to be efficient, and despite the screening step that we included, there is an important caveat to bear in mind — namely that it is important that all clones be considered only as 'candidates' until the actual abundance of their mRNA is quantitated in treated and control samples. Towards this end, we examined the expression of a limited number of clones using semi-quantitative RT-PCR. Albumin was used as the reference gene as we have previously found that the expression of this gene does not appear to change with the treatment regime that we used (Fig. 4, and data not shown). There are a number of interesting points to note from our results. The first is the presence of genes that serve as appropriate positive controls in the upregulated and downregulated series. For example, in the rat it can be seen that CYP4AI expression increases 14-fold following treatment. Although CYP4AI mRNA expression levels following WY-14,643 treatment have not been previously reported in this model, the figure compares favourably with that recorded by Bell et al. (1991), who used RNAse-protection to quantitate CYP4A1 in rat liver following treatment with methylclofenapate, another PP. In addition, we also confirmed that the peroxisomal enoyl-CoA:hydratase-3-hydroxyacyl-CoA bifunctional enzyme is also upregulated 9-fold, in agreement with the findings of Chen and Crane (1992).

A number of genes were downregulated following Wy-14,643 exposure, including CYP2C11 expression. Corton et al. (1997) reported similar findings and suggested that this may in part explain why male rats exposed to Wy-14,643 and some other PPs have high serum estradiol levels, as estradiol is a substrate for CYP2C11. We have also shown that the expression of contrapsin-like protease inhibitor (CLPI) was downregulated by Wy-14,643. This has not previously been reported, and we suggest that it may be linked to a requirement for increased availability of amino acids to accommodate the hepatomegaly induced by treatment. Although little is known of the function of parathymosin-α, ($zinc^{2+}$-binding protein) it has been shown to interact with the globular domain of histone H1, suggesting a role in histone function (Kondili et al., 1996). In contrast to the

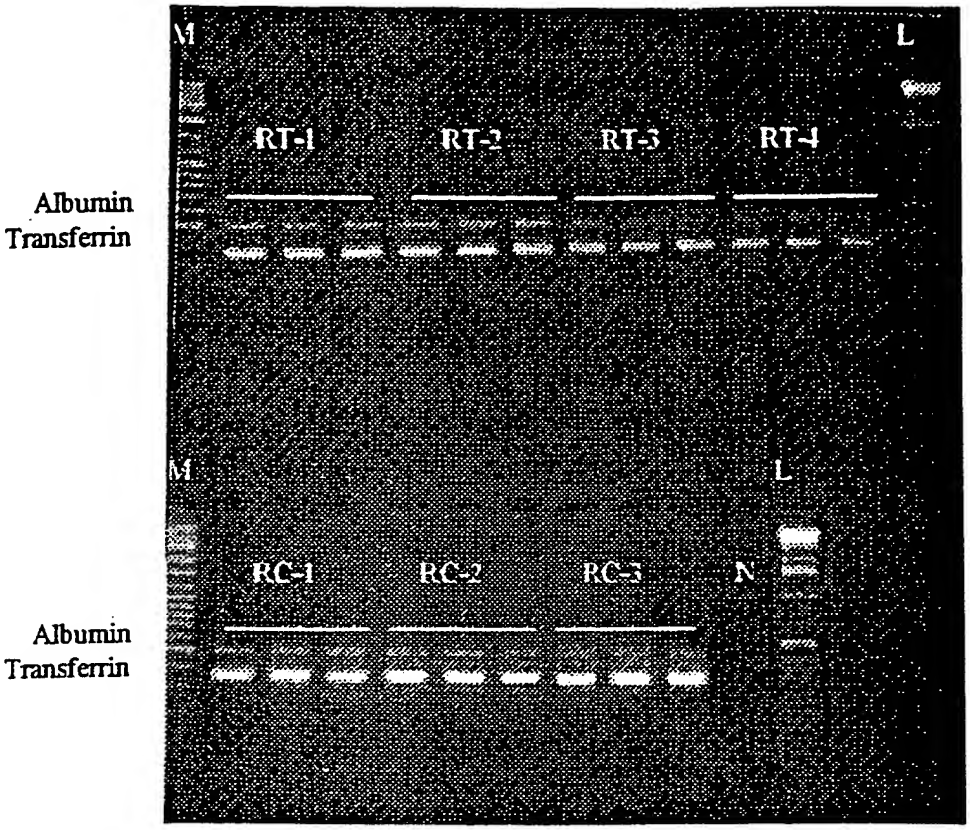


Fig. 4. Semi-quantitative RT-PCR experiment showing relative decrease in expression of transferrin in treated rat liver (RT-1 to RT-4) compared to controls (RC-1 to RC-3). An equal amount of mRNA was used in each reaction (10 ng), and each sample was quantitated in triplicate to reduce the effects of inter-tube variation. N is negative control (no mRNA). Lane M is a 100 bp ladder and lane L is a 1 Kb DNA ladder.

downregulation observed in this work, other studies have shown that parathymosin-α expression is elevated in breast cancer (Tsitsilonis et al., 1993, 1998), with the implication that parathymosin-α may somehow be involved in regulating cell proliferation by more than one mechanism. Transferrin has previously been shown to be downregulated in rat liver by hypolipidemic PPs (Hertz et al., 1996). It is therefore interesting to note that we isolated a clone identified as transferrin from the upregulated display profile. Since we confirmed by RT-PCR that transferrin is in fact downregulated in the rat (Fig. 4), we conclude that transferrin was either a false positive or was incorrectly identified. It could also be that we have isolated a close relative, splice variant or isoform of transferrin, which demonstrates a different expression profile under these experimental conditions. Further investigations are therefore required to determine which of these possibilities are correct.

One of our most intriguing observations was that one gene, CD81, appeared to be upregulated in rat liver but downregulated in guinea pig liver following Wy-14,643 exposure. CD81 is a widely expressed cell surface protein that is involved in a large number of cellular functions, including adhesion, activation, proliferation and differentiation (reviewed by Levy et al., 1998). Since all of these functions are altered to some extent in carcinogenesis, it is perhaps an important observation that CD81 expression is differentially regulated in a resistant and sensitive species exposed to a non-genotoxic carcinogen.

Albumin and ribosomal genes appear common to all differential displays and are thus undesirable false positives. However, due to their high expression in the liver, they are difficult to re-

ıove. We also noted a number of gene species, articularly in the guinea pig, which were comıon to both upregulated and downregulated rofiles. Again, the most likely reason for these aving arisen is their high abundance.

A relatively large number of upregulated and ownregulated genes were isolated from guinea ig liver following Wy-14,643 exposure. However, ıe guinea pig genome has been relatively poorly haracterised and so many of the clones were lentified as resembling genes or ESTs from other ɔecies. Without full-length sequence data it is ifficult to ascertain the accuracy of the assigned lentities and this must be borne in mind when tilising data such as this, for example, in designıg effective primers for RT-PCR studies. Alıough the actual isolated clone sequences can be sed to do this, their relatively small size often ɛstricts the ability to design effective primers. In ddition, as we observed with transferrin, using a ublished full-length sequence may help to idenfy false positives.

By comparing the expression profiles of genes showing altered expression in a PP-sensitive species (rat) with a PP-resistant species (guinea pig), it was our aim to identify genes that are mechanistically relevant to the non-genotoxic hepatocarcinogenic action of Wy-14,643. However, few of the genes that we have isolated were common to both the rat and the guinea pig. This suggests either that the molecular mechanisms of response in these two species are so different that few genes are commonly regulated in response to Wy-14,643 exposure, or that we have recovered only a small proportion of those genes that have altered expression. The latter seems the more likely scenario since it is perceived that one of the main problems of subtractive hybridisation and other differential expression technologies is the inability to consistently isolate rare gene transcripts (Bertioli et al., 1995). This is potentially problematic in that weakly expressed genes may play an important role in regulating key cellular processes, and that the majority of mRNA species are classified as

able 6
ɛmi-quantitative RT-PCR analysis of selected gene species in the rat and guinea pig[a]

ranscript	Putative change of expression following treatment according to dotblot		Change according to RT-PCR quantitation	
	Rat	Guinea pig	Rat	Guinea pig
lbumin	N/A	N/A	No change	No change
ifunctional enzyme	Up	N/A	Upregulated* (9×)	N/O
YP2C11	Down	N/A	Downregulated* (Abolished)	N/D
YP4A1	Up	N/A	Upregulated* (14×)	N/D
atalase	N/A	Up	No change	N/O
D81 (TAPA-1)	Up	Down	N/O	Upregulated**(1.4×)
ontrapsin-like protease inhibitor	Down	N/A	Downregulated** (0.5×)	N/D
ɜrathymosin-α (zinc^{2+} binding protein)	Down	N/A	Downregulated** (0.6×)	N/D
ransferrin	Up	N/A	Downregulated* (0.5×)	No change
DP-Glucuronosyl transferase	Down	N/A	Downregulated** (0.2×)	N/O
ownUnknown-1	Down	N/A	No change ($P = 0.06$)	N/D
ı-α2-glycoprotein	Up	N/A	No change	N/O

[a] N/A, not applicable; N/O, not optimised; N/D, not done.
* $P < 0.0005$;
** $P < 0.05$.

'rare' in abundance (Bertioli et al., 1995). However, in their original paper describing the SSH technique, Gurskaya et al. (1996) demonstrated that SSH can enrich rare molecules between 1000- and 5000-fold in a single round of hybridisation. Unfortunately, due to high background smearing in our initial experiments (which hindered identification of single bands), we were compelled to reduce the primary hybridisation time to only 4 h — a step that theoretically is likely to reduce the number of rare sequences (CLONTECHniques, 1996). Furthermore, it has been claimed by the manufacturers that, whilst this technique can identify changes as small as 1.5-fold between the driver and tester populations, it is best suited to the isolation of genes that show a greater than 5-fold increase (CLONTECHniques, 1996). In addition, where tester and driver contain genes with large and small differences in abundance, the SSH method will be biased towards identifying those genes with the large differences (CLONTECHniques, 1996). Thus, it is most probable that we have not isolated all of the more rarely expressed transcripts and those demonstrating small changes in expression.

One problem that remains is identifying the function of genes isolated in SSH experiments as described herein, some of which may be crucial to the process of carcinogenesis, and are, to date, unidentified. However, we have provided evidence herein that SSH can be used to begin the process of characterising the extent and importance of altered gene expression in response to a chemical stimulus. The developments of this approach should include characterisation of temporal and dose responses, and functional analysis studies including knockout mice. In combination, such studies should make a significant contribution to our understanding of the molecular mechanisms of action and physiological relevance of gene regulation in non-genotoxic hepatocarcinogenesis. It should then be possible to ascertain whether differentially expressed genes are causally or casually related to the chemical-induced toxicity, and therefore a substantial mechanistic advance.

It is clear that there are also broader applications for this experimental approach that go beyond understanding the molecular mechanisms of peroxisome-proliferator induced non-genotoxic hepatocarcinogenesis in rodents. The potential medical and therapeutic benefits of elucidating the molecular changes that occur in any given cell in progressing from the normal to the carcinogenic (or other diseased, abnormal or developmental) state are very substantial. Notwithstanding the lack of complete functional identification of altered gene expression, such gene profiling studies described herein essentially provides a 'fingerprint' of each stage of carcinogenesis, and should help in the elucidation of specific and sensitive biomarkers for different types of cancer. Amongst other benefits, such fingerprints and biomarkers could help uncover differences in histologically identical cancers, and provide diagnostic tests for the earliest stages of neoplasia. In addition, the genes identified by this approach may be incorporated into gene-chip DNA-arrays, thus providing a standard genetic fingerprint for a particular toxin treatment in a particular species. Interrogation of these gene arrays for an unknown compound that has a similar pattern to the known reference chemical would then provide evidence that the unknown may have a toxicity profile similar to the 'standard' fingerprint, thereby serving as a mechanistically relevant platform for further detailed investigations.

Acknowledgements

This work was funded by Rhone-Poulenc Agrochemicals, Nice, France.

References

Anderson, N.L., Esquer-Blasco, R., Richardson, F., Foxworthy, P., Eacho, P., 1996. The effects of peroxisome proliferators on protein abundances in mouse liver. Toxicol. Appl. Pharmacol. 137, 75–89.

Ashby, J., 1992. Prediction of non-genotoxic carcinogenesis. Toxicol. Lett. 64–65, 605–612.

Bell, D.R., Bars, R.G., Gibson, G.G., Elcombe, C.R., 1991. Localisation and differential induction of cytochrome P450IVA and acyl coA oxidase in rat liver. Biochem. J. 275, 247–252.

Bertioli, D.J., Schlichter, U.H.A., Adams, M.J., Burrows, P.R., Steinbiss, H.-H., Antoniw, J.F., 1995. An analysis of

differential display shows a strong bias towards high copy number mRNAs. Nucleic Acid Res. 23 (21), 4520–4523.

Cattley, R.C., Kato, M., Popp, J.A., Teets, V.J., Voss, K.S., 1994. Initiator-specific promotion of hepatocarcinogenesis by Wy-14,643 and clofibrate. Carcinogenesis 15 (8), 1763–1766.

Chen, N., Crane, D.I., 1992. Induction of the major integral membrane protein of mouse liver peroxisomes by peroxisome proliferators. Biochem J. 283, 605–610.

CLONTECHniques, 1996. Technical Tips: Clontech PCR-Select cDNA Subtraction, October 25, application notes.

CLONTECHniques, 1997. PCR-Select Differential Screening Kit — The Next Step After Clontech PCR-Select cDNA Subtraction. XII(2), 18–19, application notes.

Corton, J.C., Bocos, C., Moreno, E.S., Merrit, A., Cattley, R.C., Gustaffson, J.A., 1997. Peroxisome proliferators alter the expression of estrogen-metabolising enzymes. Biochimie 79, 151–162.

Davis, M., Cohen, D.I., Nielson, E.A., Steinmetz, M., Paul, W.E., Hood, L., 1984. Cell-type-specific cDNA probes and the murine I region: the localisation and orientation of Ad/a. Proc. Natl. Acad. Sci USA 81, 2194–2198.

Diatchenko, L., Lau, Y.-F.C., Campbell, A.P., Chenchik, A., Moqadam, F., Huang, B., Lukyanov, K., Gurskaya, N., Sverdlov, E.D., Siebert, P.D., 1996. Suppression subtractive hybridisation: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. Proc. Natl. Acad. Sci. USA 93, 6025–6030.

Duguid, J., Dinauer, M., 1990. Library subtraction of in vitro cDNA libraries to identify differentially expressed genes in scrapie infection. Nucleic Acid Res. 18 (9), 2789–2792.

Geisinger, A., Rodriguez, R., Romero, V., Wettstein, R., 1997. A simple method for screening cDNAs arising from the cloning of RNA differential display bands. Elsevier trends journals technical tips online, http://tto.trends.com, document number T01110

Guimaraes, M.J., Lee, F., Zlotnik, A., McClanahan, T., 1995. Differential display by PCR: novel findings and applications. Nucleic Acid Res. 23 (10), 1832–1833.

Gurskaya, N.G., Diatchenko, L., Chenchik, P.D., Siebert, P.D., Khaspekov, G.L., Lukyanov, K.A., Vagner, L.L., Ermolaeva, O.D., Lukyanov, S.A., Sverdlov, E.D., 1996. Equalising cDNA subtraction based on selective suppression of polymerase chain reaction: cloning of Jurkat cell transcripts induced by phytohemaglutinin and phorbol 12-myrystate 13-acetate. Anal. Biochem. 240, 90–97.

Hayashi, F., Tamura, H., Yamada, J., Kasai, H., Suga, T., 1994. Characteristics of the hepatocarcinogenesis caused by dehydroepiandrosterone, a peroxisome proliferator, in male F-344 rats. Carcinogenesis 15 (190), 2215–2219.

Hedrick, S.M., Cohen, D.I., Nielsen, E.A., Davis, M.M., 1984. Isolation of cDNA clones encoding T cell-specific membrane-associated proteins. Nature 308 (8), 149–153.

Hertz, R., Seckbach, M., Zakin, M.M., Bar-Tana, J., 1996. Transcriptional suppression of the transferin gene by hypolipidemic peroxisome proliferators. J. Biol. Chem. 271 (1), 218–224.

Hubank, M., Schatz, D.G., 1994. Identifying differences in mRNA expression by representational difference analysis. Nucleic Acid Res. 22 (25), 5640–5648.

Human and Experimental Toxicology, 1994. Hum. Exp. Toxicol. 13 (Suppl. 2) (entire issue).

Kolls, J., Dsininger, P., Cohen, C., Larson, J., 1993. cDNA equalisation for reverse transcription-polymerase chain reaction quantitation. Anal. Biochem 208, 264–269.

Kondili, K., Tsolas, O., Papamarcaki, T., 1996. Selective interaction between parathymosin and histone H1. Eur. J. Biochem. 242 (1), 67–74.

Lake, B.G., Evans, J.G., Gray, T.J.B., Korosi, S.A., North, C.J., 1989. Comparative studies of nafenopin-induced hepatic peroxisome proliferation in the rat, Syrian hamster, guiea pig and marmoset. Toxicol. Appl. Pharmacol. 99, 148–160.

Lake, B.G., Evans, J.G., Cunninghame, M.E., Price, R.J., 1993. Comparison of the hepatic effects of Wy-14,643 on peroxisome proliferation and cell replication in the rat and Syrian hamster. Environ. Health Perspect. 101 (S5), 241–248.

Levy, S., Todd, S.C., Maecker, H.T., 1998. CD81 (TAPA-1): a molecule involved in signal transduction and cell adhesion in the immune system. Annu. Rev. Immunol. 16, 89–109.

Li, W.B., Gruber, C.E., Lin, J.J., D'Alessio, J.M., Jessee, J.A., 1994. The isolation of differentially expressed genes in fibroblast growth factor stimulated BC3H1 cells by subtractive hybridization. BioTechniques 16, 722–729.

Liang, P., Pardee, A.B., 1992. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. Science 257 (5072), 967–971.

Liang, P., Averboukh, L., Keyomarsi, K., Sager, R., Pardee, A.B., 1992. Differential display and cloning of messenger RNAs from human breast cancer versus mammary epithelium. Cancer Res. 52, 6966–6968.

Lipman, D.J., Pearson, W.R., 1985. Rapid and sensitive protein similarity searches. Science 227, 1435–1441.

Makowska, J.M., Gibson, G.G., Bonner, F.W., 1992. Species differences in ciprofibrate induction of hepaic cytochrome P450IVA1 and peroxisome proliferation. J. Biochem. Toxicol. 7, 183–191.

Marsman, D.S., Cattley, R.C., Conway, J.G., Popp, J.A., 1988. Relationship of hepatic peroxisome proliferation and replicative DNA synthesis to the hepatocarcinogenicity of the peroxisome proliferators di-(2-ethylhexyl)phthalate and [4-chloro-6-(2,3-xylidino)-2-pyrimidinylthio]acetic acid (Wy-14,643) in rats. Cancer Res. 48, 6739–6744.

Mathieu-Daude, F., Cheng, R., Welsh, J., McClelland, M., 1996. Screening of differentially amplified cDNA products from RNA arbitrarily primed PCR fingerprints using single strand conformation polymorphism (SSCP) gels. Nucleic Acid Res. 24 (8), 1504–1507.

Orton, T.C., Adam, H.K., Bentley, M., Holloway, B., Tucker, M.J., 1984. Clobuzarit: species differences in the morphological and biochemical response of the liver following chronic administration. Toxicol. Appl. Pharmacol. 73, 138–151.

Parodi, S., 1992. Non-genotoxic factors in the carcinogenic process: problems of detection and hazard evaluation. Toxicol. Lett. 64–65, 621–630.

Pearson, W.R., Lipman, D.J., 1988. Imported tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA 85, 2444–2448.

Rockett, J.C., Esdaile, D.J., Gibson, G.G., 1997. Molecular profiling of non-genotoxic hepatocarcinogenesis using differential display reverse transcription-polymerase chain reaction (ddRT-PCR). Eur. J. Drug. Metab. Pharmacokinet 22 (4), 329–333.

Rodricks, J.V., Turnbull, D., 1987. Inter-species differences in peroxisomes and peroxisome proliferation. Toxicol. Ind. Health 3, 197–212.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. In: Ford, N., Nolan, C., Ferguson, M. (Eds.), Molecular Cloning — A Laboratory Manual, second ed. Cold Spring Harbor Laboratory Press, New York.

Sargent, T., Dawid, I., 1983. Differential gene expression in the gastrula of *Xenopus laevis*. Science 222, 135–139.

Smith, N.R., Li, A., Aldersley, M., High, A.s., Markham, A.F., Robinson, P.A., 1997. Rapid determination of the complexity of cDNA bands extracted from DDRT-PCR polyacrylamide gels. Nucleic Acid Res. 25 (17), 3552–3554.

Tsitsilonis, O.E., Stiakakis, J., Koutselinis, A., Gogas, J., Markopoulos, C., Yialouris, P., Bekris, S., Panoussopoulos, D., Kiortsis, V., Voelter, W., Haritos, A.A., 1993. Expression of alpha-thymosins in human tissues in normal and abnormal growth. Proc. Natl. Acad. Sci. USA 90 (20), 9504–9507.

Tsitsilonis, O.E., Bekris, E., Voutsas, I.F., Baxevanis, C.N., Markopoulos, C., Papadopoulou, S.A., Kontzoglou, K., Stoeva, S., Gogas, J., Voelter, W., Papamichail, M., 1998. The prognostic value of alpha-thymosins in breast cancer. Anticancer Res. 18 (3A), 1501–1508.

Wan, J.S., Sharp, S.J., Poirier, G.M.-C., Wagaman, P.C., Chambers, J., Pyati, J., Horn, Y.-L., Galindo, J.E., Huvar, A., Peterson, P.A., Jackson, M.R., Erlander, M.G., 1996. Cloning differentially expressed mRNAs. Nat. Biotechnol. 14, 1685–1691.

Wawer, C., Ruggeberg, H., Meyer, G., Muyzer, G., 1995. A simple and rapid electrophoresis method to detect sequence variation in PCR-amplified DNA fragments. Nucleic Acid Res. 23 (23), 4928–4929.

TOXICOLOGY

An international journal concerned with the effects of chemicals on living systems and immunotoxicology

ELSEVIER

Special Issue

Festschrift dedicated to Professor Dr. K.J. Netter

ELSEVIER

Toxicology 144 (2000) 13–29

www.elsevier.com/locate/toxicol

Use of suppression-PCR subtractive hybridisation to identify genes that demonstrate altered expression in male rat and guinea pig livers following exposure to Wy-14,643, a peroxisome proliferator and non-genotoxic hepatocarcinogen

John C. Rockett [1], Karen E. Swales, David J. Esdaile [2], G. Gordon Gibson *

Molecular Toxicology Group, School of Biological Sciences, University of Surrey, Guildford, Surrey GU2 5XH, UK

Abstract

Understanding the genetic profile of a cell at all stages of normal and carcinogenic development should provide an essential aid to developing new strategies for the prevention, early detection, diagnosis and treatment of cancers. We have attempted to identify some of the genes that may be involved in peroxisome-proliferator (PP)-induced non-genotoxic hepatocarcinogenesis using suppression PCR subtractive hybridisation (SSH). Wistar rats (male) were chosen as a representative susceptible species and Duncan–Hartley guinea pigs (male) as a resistant species to the hepatocarcinogenic effects of the PP, [4-chloro-6-(2,3-xylidino)-2-pyrimidinylthio] acetic acid (Wy-14,643). In each case, groups of four test animals were administered a single dose of Wy-14,643 (250 mg/kg per day in corn oil) by gastric intubation for 3 consecutive days. The control animals received corn oil only. On the fourth day the animals were killed and liver mRNA extracted. SSH was carried out using mRNA extracted from the rat and guinea pig livers, and used to isolate genes that were up and downregulated following Wy-14,643 treatment. These genes included some predictable (and hence positive control) species such as CYP4A1 and CYP2C11 (upregulated and downregulated in rat liver, respectively). Several genes that may be implicated in hepatocarcinogenesis have also been identified, as have some unidentified species. This work thus provides a starting point for developing a molecular profile of the early effects of a non-genotoxic carcinogen in sensitive and resistant species that could ultimately lead to a short-term assay for this type of toxicity.

Keywords: Wy-14,643; Peroxisome proliferator; Non-genotoxic hepatocarcinogenesis; Suppression PCR subtractive hybridisation; RT-PCR; Rat; Guinea pig; Gene regulation; Differential gene display; Gene profiling

* Corresponding author. Tel.: +44-1483-259704; fax: +44-1483-576978.
E-mail address: g.gibson@surrey.ac.uk (G.G. Gibson)
[1] Present address: US Environmental Protection Agency, National Health and Environmental Effects Research Laboratory, Reproductive Toxicology Section, Research Triangle Park, NC 27711, USA.
[2] Present address: Rhone-Poulenc Agrochemicals, Toxicology Department, Sophia-Antipolis, Nice, France.

0300-483X/00/$ - see front matter
PII: S0300-483X(99)00214-0

Introduction

Γhe advent of combinatorial chemistry and nputer-aided drug design has led to a recent ɜurge in the number of chemical compounds ıt have potential therapeutic, agricultural and lustrial applications. Although it has been sugted that the contribution of synthetic chemicals the overall incidence of human cancer is low, re still remains an absolute requirement to ıluate all new chemicals for toxic and carcinoıic potential. The latter is one of the most ›blematic areas of chemical safety evaluation l is usually carried out using short-term in vitro l in vivo genotoxicity assays augmented by onic bioassay tests. The short-term assays have ›ved useful in the early identification of poten- genotoxic carcinogens, but their value is limI by observations that suggest that ›roximately 60% of chemicals identified as carɔgens in life-exposure studies produce mainly ative findings in short-term genotoxcity tests hby, 1992; Parodi, 1992). Thus, there is curtly no reliable and rapid means of evaluating carcinogenic risk of new chemicals that fall › this latter group of compounds, termed nonotoxic (or epigenetic) carcinogens.

)ne approach to addressing this problem is to :idate the molecular mechanisms by which ›wn non-genotoxic carcinogens act. It should ı be possible to identify common factors/ :hanisms that can serve as early biomarkers of :inogenic potential for new chemicals. To this , a large number of groups have reported on various effects of non-genotoxic compounds various animal species (Marsman et al., 1988; :e et al., 1993; Cattley et al., 1994; Hayashi et 1994; Human and Experimental Toxicology, 4; Anderson et al., 1996). However, the mechtic picture is still far from complete with many hose genes involved in the carcinogenic proremaining unknown, and their identification efore remains a key goal in elucidating the ɜcular mechanisms by which non-genotoxic inogenesis occurs.

ıbtractive hybridisation (SH) and related tech›gies such as representational difference analy- (RDA) (Hubank and Schatz, 1994) and differential display (DD) (Liang and Pardee, 1992) can be used to aid the isolation of genes showing altered expression in target tissues following exposure to a chemical stimulus. These techniques can also be used to identify differential gene expression in neoplastic and normal cells (Liang et al., 1992), infected and normal cells (Duguid and Dinauer, 1990), differentiated and undifferentiated cells (Sargent and Dawid, 1983; Guimaraes et al., 1995), activated and dormant cells (Gurskaya et al., 1996; Wan et al., 1996), different cell types (Hedrick et al., 1984; Davis et al., 1984) amongst others. Most importantly, using such approaches, no prior knowledge of the specific genes that are upregulated/downregulated is required.

Using a variation of SH, termed suppression-PCR subtractive hybridisation (SSH) (Diatchenko et al., 1996), we have previously reported the isolation of a number of genes showing altered expression in male rat liver following acute exposure to phenobarbital (Rockett et al., 1997). In the current work we have used the same experimental approach to isolate genes that are differentially expressed in the livers of male rats and guinea pigs following short-term (3-day) exposure to the peroxisome proliferator (PP) and non-genotoxic hepatocarcinogen, Wy-14,643. We have isolated and identified a number of gene species, some of which may be important in the induction of, or protection against, non-genotoxic hepatocarcinogenesis.

2. Materials and methods

2.1. Animals and treatment

All animal experiments were undertaken in accordance with Her Majesty's Home Office Department guidelines under the auspices of approved personal and project licences. Male Wistar rats (150–200 g) and male Duncan–Hartley guinea pigs (250–300 g) were obtained from Kingman and Bantam (Hull, UK). Upon receipt, both groups were randomly assigned into two groups of four. They were maintained on a rat, mouse or guinea pig standard diet (B&K Univer-

sal, Hull) and a daily cycle of alternating 12-h periods of dark and light. The room temperature was maintained at 19°C and a relative humidity of 55%. The animals were acclimatised to this environment for 7 days before treatment commenced. [4-chloro-6-(2,3-xylidino)-2-pyrimidinylthio] acetic acid (Wy-14,643, Campo, Emmerich; 250 mg/kg per day in corn oil) was administered by gavage to the treated groups of rats and guinea pigs on 3 consecutive days, whilst control groups received an equal volume of corn oil only. During this time, all animals had free access to food and water. The animals were killed by cervical dislocation on the fourth day, and their livers immediately excised, weighed, sliced into approximately 0.5-cm cubes, snap frozen in liquid nitrogen and stored at -70°C.

2.2. mRNA extraction

Approximately 0.25 g of each frozen liver sample was ground under liquid nitrogen using a mortar and pestle. Messenger RNA was extracted from the ground liver using the PolyATtract® System 1000 kit (Promega, Madison, USA) according to the technical manual provided by the manufacturers. The mRNA was DNase-treated (RQ Rnase-free Dnase, Promega, final concentration 10 U/ml) before phenol/chloroform extraction and ethanol precipitation. The mRNA was redissolved at a final concentration 500–1000 ng/μl.

2.3. cDNA Subtraction

This was carried out using the PCR-Select™ cDNA Subtraction Kit (Clontech, Palo Alto, USA) according to the manufacturer's instructions. Subtractions were carried out with mRNAs derived from single animals. The mRNA from the remaining three animals in each group was later used for quantitative RT-PCR analysis of specific genes.

2.4. Band extraction and cloning

The secondary PCR reactions from the cDNA subtraction procedure were run on a 2% Metaphor agarose gel (FMC, Rockland, USA) containing 0.5 μg/ml ethidium bromide (Sigma, Dorset, UK). One times TAE (0.04 M Tris–acetate, 0.001 M EDTA) was used to prepare the gel and as the running buffer. After running for 6–7 h at 3.75 V/cm, the gel was overstained for 30 min with SYBR Green I DNA stain (FMC, 1:10 000 dilution in 1 × TAE). Each discernible band of the differential display pattern was extracted from the gel with a scalpel and the DNA eluted using a Genelute™ agarose spin column (Supelco, Bellefonte, USA). Five microlitres of the eluted DNA was reamplified using the original nested (secondary) PCR primers supplied with the PCR-Select™ cDNA subtraction kit. The PCR products were electrophoresed on a 2% standard agarose gel (Boehringer Mannheim, East Sussex, UK) and the reamplified target bands extracted from the gel as above. The eluted DNA was immediately ligated into a TOPO TA Cloning® vector (Invitrogen, Carlsbad, USA) before transformation in *Escherichia coli* TOP10F' One Shot™ cells (Invitrogen).

2.5. Colony screening

2.5.1. Stage I

Eight transformed (white) colonies from each band were grown up for 6 h in 200 μl LB broth containing ampicillin (Sigma, 50 mg/ml). One microlitre of this was subjected to PCR using the same conditions and nested primers as described above. One tenth (2 μl) of the completed PCR reaction was electrophoresed on a 2% standard agarose gel and one tenth on a 2% standard agarose gel containing HA red (Hanse Analytik GmbH, Bremen, Germany, 1 U/ml) to discern the differentially cloned products. The remainder of the PCR reaction was used to prepare duplicate dotblots on Hybond N$^+$ membranes (Amersham, Little Chalfont, UK).

2.5.2. Stage II

The duplicate dotblots were probed with (a) the final differential display reaction and (b) the 'reverse-subtracted' differential display reaction. To make the 'reverse-subtracted' probe, the subtractive hybridisation step of the differential display

.T-PCR procedure was carried out using the riginal tester (treated) mRNA as the driver and ıe original driver (control) mRNA as the tester. robing and visualisation were carried out using ıe ECL direct nucleic acid labelling and detecon system (Amersham, Little Chalfont, UK) acɔrding to the manufacturer's instructions. Those .ones that were positive for (a) but negative for ɔ), or showed a substantially larger positive sigal with (a) compared to (b), were selected for NA sequence analysis.

.6. *DNA sequencing*

The remainder of the cultures (prepared in age I screening) containing different cloning roducts (as discerned in the two screening steps) ere grown up overnight in 5 ml LB broth con.ining ampicillin (50 mg/ml). A plasmid miniprep as prepared from each (Wizard Plus SV linipreps DNA purification system, Promega) :cording to the manufacturer's instructions. The oned inserts were sequenced on an automated BI DNA sequencer (Applied Biosystems, Warngton, UK) using the M13 forward primer ƷTAAAACGACGGCCAGT) or M13 reverse imer (AACAGCTATGACCATG).

7. *Identification of differentially regulated genes*

Gene sequences thus obtained were identified ;ing the FASTA 3.0 programme (Lipman and ;arson, 1985; Pearson and Lipman, 1988) (http:/ /ww.ddbj.nig.ac.jp/E-mail/homology.html) to arch all EMBL databases for matching DNA quences. Each clone sequence was submitted in e forward and reverse direction, and the one turning the highest statistical probability of atch to a known sequence was noted. Sequence ɔmologies between our submitted clone sequence ıd the queried database sequence were deterined (by FASTA) over a region of at least 60 ıse pairs.

8. *RT-PCR analysis of selected candidate genes*

cDNA sequences of the target genes were obined from the NIH gene database (GenBank at http://www.ncbi.nlm.nih.gov/Web/Search/index.html) and the computer programme GENE JOCKEY (BioSoft, Cambridge, UK) used to select primer pairs from these sequences. Where guinea pig sequences were available, rat and guinea pig sequences were aligned and primers chosen from regions of homology. If guinea pig sequences were not available, rat and human sequences were used. In cases where exact homology could not be found, the sequence from the rat was used. In the case of CD81 only, no rat or guinea pig sequences were available and so mouse and human sequences were aligned and a primer pair chosen from a region of homology. Primers (obtained from Gibco-BRL, Paisley, UK) were dissolved at a concentration of 50 pmol/μl in sterile distilled water and stored at $-20°C$. The primer pairs used plus other reaction parameters are shown in Table 1. mRNA was extracted (as described above) from all four treated animals and from three animals in the control group. Integrity of the eluted mRNA was confirmed on a 2% agarose gel, and the concentration and purity were measured using a Genequant II spectrophotometer (LKB, Bromma, Sweden) and then diluted to 10 ng/μl. One microlitre of this latter solution was used per RT-PCR reaction.

RT-PCR was carried out in a single tube (50 μl) reaction using the Access RT-PCR system (Promega) according to manufacturer's instructions. In the kinetic and quantitative analyses, omission of RNA was used as a control for the presence of any contaminating DNA. After obtaining a PCR signal of the correct size and optimising the reaction conditions, each PCR product was digested with between two and four separate restriction enzymes. Specific restriction patterns were thus obtained, which further confirmed the identity of the PCR products as being the original target genes. Kinetic analysis (14–32 cycles) was then performed in each case to determine the location of the mid-log phase.

For the semi-quantitative analysis of each target gene, RT-PCR reactions were carried out in triplicate for each sample to reduce the effect of intertube RT-reaction variations (Kolls et al., 1993) and pipetting errors. For each gene, a mastermix containing enough reagents for three times

Table 1
Primer sequences and reaction conditions used in semi-quantitative RT-PCR analysis of selected genes

Transcript	Genbank accession No.	Primer sequences		Size of rat PCR product (bp)	Annealing temperature (°C) rat/guinea pig	No. of PCR cycles rat/guinea pig
		upstream	downstream			
Albumin	J00698 (rat)	TGGAGAGA-GAGC-CTTCAAAGC	CTTAG-CAAGTCTCAGCAGCA	436	60/59	15/22
Bifunctional enzyme	K03249 (rat)	GCACC-CACTTCTTCT-CACCAGC	TGGCAATGATG-GTCCAGTAAGG	347	57/–	21/–
CYP2C11	J02657 (rat)	CCATCATGACC-CTGAGG	GAAGTCCCGAG-GATTGT	410	50/–	20/–
CYP4A1	M14972 (rat)	GATGGCTGCAC-CATGAG	GGCCTTTG-GATCTGATC	357	57/–	22/–
Catalase	M11670 (rat)	ACCAAATACTC-CAAGGCAAAGG	GCCCTG-GTCAGTCTTG-TAATGG	450	63/–	27/–
CD81 (TAPA-1)	X59047 (mouse)	ATTTCGTCTTCTGGCTGGCTGG	GCCTGGTCATA-GAACTGCTTCA	337	57/59	23/22
Contrapsin-like protease inhibitor	RNCCP23 (rat)	GACTATGTGAG-CAATCAGAC	GTCTCTGGTTG-CAAGCT	341	50/–	20/–
Parathymosin-α (Zn^{2+} binding protein)	X64053 (rat)	CGGCACCAT-GTCGGAGAAGA	TTGTGTGTTCCT-GCCCCACC	382	62/–	24/–
Transferrin	D38380 (rat)	AGCTGTGT-CAACTGT-GTCCAGG	GAGGAGAGCC-GAACAGTTG-GAA	360	57/59	22/22
UDP-GT	U06273 (rat)	GGAT-GTCTGGGAAGTGG	GCAGTTCAGC-TATCAGCT	495	50/–	23/–
DownUnknown-1	n/a	CGACGTTTC-CAAGGCA	TGTTGCGGCA-GAGTGGA	318	55/–	25/–
Znα2glycoprotein	D21058 (rat)	CAAATAACA-GAAGCAGTG-GAGC	GACTTCCAC-CTCCATCCAGG	433	57/–	23/–

the number of samples (seven for rat, six for guinea pig) was prepared except that mRNA was omitted, the latter being added after aliquoting 49 μl of the mastermix into an appropriate number of tubes. Amplification of albumin (the reference gene) was carried out in separate tubes since the mid-log phase of this gene is at a much lower cycle number than the target genes due to its high abundance. All RT-PCR products were analysed on 2% agarose gels containing 0.5 μg/ml ethidium bromide. The target gene samples were loaded on the gel first and run in at 3 V/cm for 10 min. The corresponding albumin samples were then loaded and the gel run for a further 1/2 h. In this way, all RT-PCR products from each target gene and albumin from the corresponding samples could be run on the same gel. Gels were photographed using type 665 posi-neg film (Sigma) and quantitation of the band intensity was carried out using a dual wavelength flying spot laser scanner densitometer (Shimadzu).

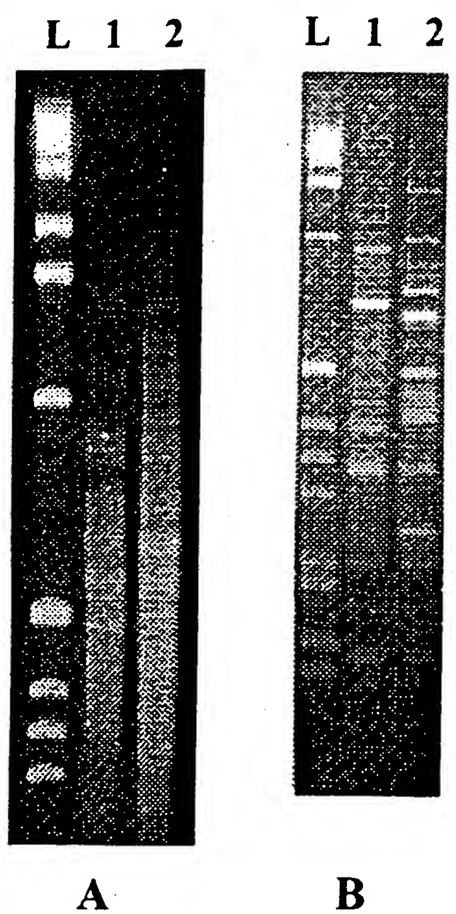


Fig. 1. Final displays of differentially expressed genes that were (1) upregulated and (2) downregulated in rat (A) and guinea pig (B) livers following 3-day treatment with Wy-14,643. mRNA extracted from control and treated livers was used to generate the differential displays using the PCR-Select cDNA subtraction kit (Clontech). Lane (L) is a 1 Kb DNA Ladder standard and 10 μl of secondary PCR reaction were loaded in all other lanes.

2.9. Statistical analysis

Statistical analysis of unpaired samples was carried out using the two-tailed Student's t-test. Values were considered statistically significant at $P < 0.05$ or less.

3. Results

3.1. Cloning and screening of transcripts

For both the rat and guinea pig experimental groups, cDNA subtraction was carried out in the forward (control driving tester) and reverse (tester driving control) directions to isolate both upregulated and downregulated mRNA species respectively. Using a standard primary hybridisation time of 8 h we obtained a substantial amount of non-specific products in all the final differential displays (data not shown). This background smearing was almost completely removed by reducing the primary hybridisation time to 4 h (CLONTECHniques, 1996). Fig. 1 shows the ddRT-PCR patterns of genes showing altered expression in rat and guinea pig liver following 3-day treatment with Wy-14,643. The profiles are unique for each species, and in each case the profile for the upregulated genes (control mRNA driving tester mRNA) is different to that obtained for the downregulated genes (tester mRNA driving control mRNA).

The practical outcome of the SSH method is that a series of differentially expressed genes is observed as a ladder on an agarose gel. The majority of these gene fragments fall within the 150–2000 bp range, with bands up to 5 Kbp occasionally being observed. Each band may theoretically consist of one or more products of similar size, as the gel has a maximum resolution

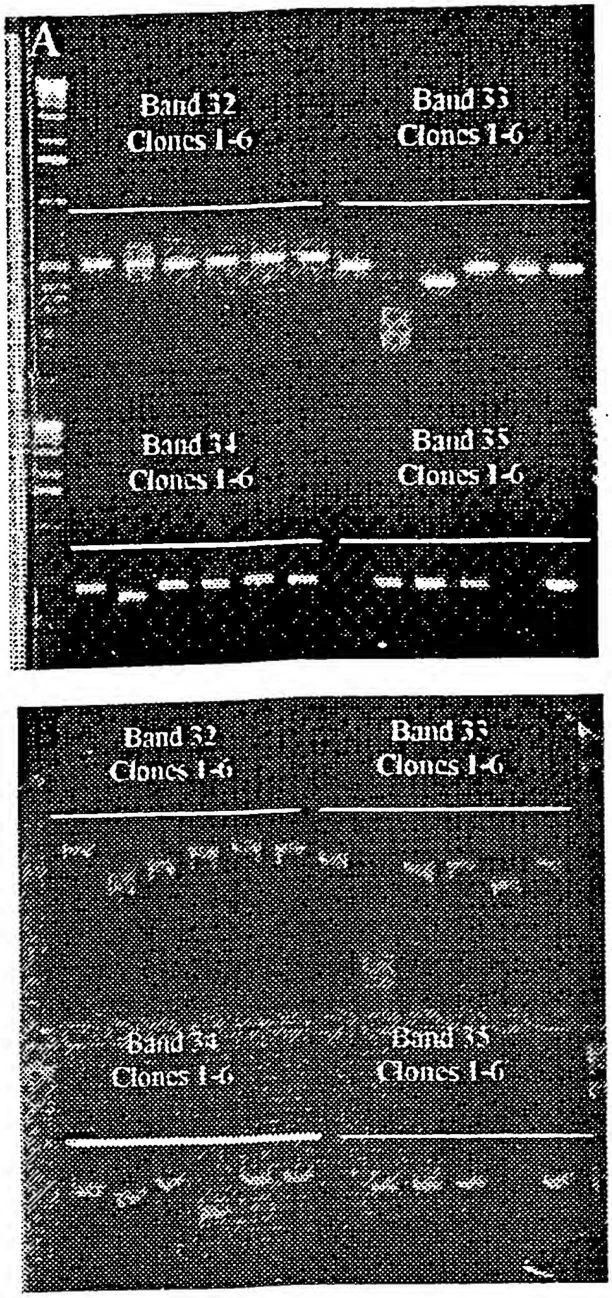


Fig. 2. Discrimination of different ddRT-PCR products having the same molecular size using HA-red. Gel (A) is a 2% standard agarose gel. Gel (B) is a 2% standard agarose gel containing 1 U/ml HA-red. Band numbers refer to the sequential bands (largest to smallest) extracted from the original display of genes upregulated in rat liver following 3-day treatment with Wy-14,643. Ten micorlitres of each PCR reaction were loaded per lane.

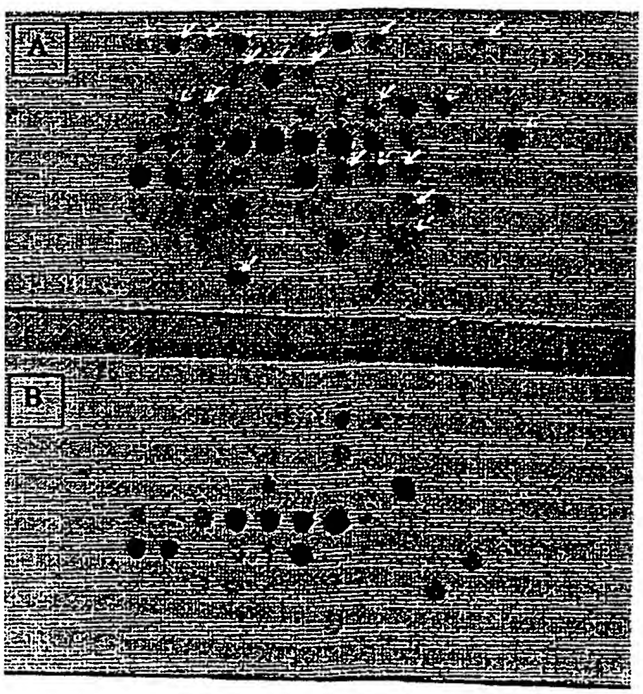
of approximately 1.5% (3 bp per 200). In addition, there may be two or more products that are the same size, but have a different sequence. Therefore some form of discrimination must be employed to isolate as many of these products as possible. HA-red screening (Geisinger et al., 1997) of a number of clones derived from each band provided a means to discriminate between different gene species of the same size. A typical example of such a gel is shown in Fig. 2. In total, 88 and 48 apparently different clones were obtained from the final differential expression patterns of upregulated and downregulated rat genes, respectively. Sixty nine and 89 apparently different clones were obtained from the final differential expression patterns of the upregulated and downregulated guinea pig genes, respectively.

Having identified as many different candidate gene products as possible in the screening step I, a second screening step was carried out on every clone to confirm those that represented true differentially expressed genes. This is necessary since no subtraction technique is 100% efficient. The approach we used, termed PCR-select differential screening (as recommended in Clontech's PCR-select cDNA subtraction kit protocol), utilises the forward and reverse subtractions as an aid to screening for the true differentially expressed genes (CLONTECHniques, 1997). Because these probes have already undergone subtraction, they have been enriched for differentially expressed genes and are therefore more sensitive than unsubtracted driver/tester cDNA probes for detecting true differential expression. All the clones that were isolated from each display were dotblotted and probed with the display from which they was obtained, plus the corresponding reverse-subtracted display. An example of such a blot is shown in Fig. 3. Clones corresponding to authentic differentially expressed mRNAs hybridised with the subtracted cDNA probe, but not the reverse-subtracted probe. We also included in the authentic positives, those clones that gave a substantially greater signal with the subtracted probe compared to the reverse-subtracted probe. False positives hybridised with either both probes or with neither probe. Of the original 88 upregulated and 48 downregulated rat clones selected for this screening step, 28 (32%) and 15 (31%) respectively, were found to be true positives. In the rat,

3 (100%) of the true positive upregulated genes 'able 2) and 11 (73%) of the true positive down- gulated genes (Table 3) were non-redundant. Of .e original 69 upregulated and 89 downregulated iinea pig clones selected for this screening step, 3 (70%) and 37 (42%) respectively, were found to : true positives. Thirty six (75%) of the upregu- ted genes (Table 4) and 33 (89%) of the down- gulated genes (Table 5) were non-redundant.

2. Identification of clones

On sequence analysis it was found that some ones were unsequencable in the first instance 113 forward primer) due to long polyA runs at appeared to prematurely terminate the se- iencing reaction. These clones were therefore sequenced from the opposite direction using the 13 reverse primer. Those xenobiotic-modulated ne products identified to date are listed in Ta- es 2 and 3 (rat) and Tables 4 and 5 (guinea pig).



. 3. Dot blots of clones of putative upregulated gene species ated from guinea pig liver following 3-day treatment with -14,643. All clones identified in the stage I screening step methods) were blotted and probed with (A) the differen- display from which they originated (control driving .ted) and (B) the reverse subtraction (treated driving con-). Arrows indicate some of the true differentially expressed ies.

Table 2
Identification of genes that were upregulated in male rat liver following 3-day treatment with WY-14,643

FASTA-EMBL gene identification (rat unless otherwise stated)	Accession No.	Sequence homology[a] (%)
Carnitine octanoyl transferase	RN26033	99
NCI_CGAP_Li1 (*H. sapiens*) (EST[b])	HS1275949	98
Peroxisomal enoyl hydratase-like protein	RN08976	98
Liver fatty acid binding protein	V01235	96
Soares mouse p3NMF19.5 *M. musculus* cDNA clone	AA038051	96
Cytochrome p450IVA1	RNCYPLA	94
Mit. 3-hydroxyl-3-methylglutaryl CoA synthase	RNHMGCOA	94
Rabgeranylgeranyl transferase component B	RNRABGERA	94
Genes for 18S, 5.8S, and 28S ribosomal RNAs	RNRRNA	94
Carnitine acetyl transferase (mouse)	MMRNACAR	92
Soares mouse NML (EST)	MM1157113	92
Bone marrow stromal fibroblast (*H. sapiens*) cDNA clone HBMSF2E4 (EST)	AA545726	92
7.5dpc embryo (mouse) (EST)	AA408192	92
Alpha-1-macroglobulin	RNALPH1M	91
Transferrin	RNTRANSA	91
Lecithin:cholesterol acyltransferase	RNU62803	90
Zn-α2-glycoprotein	RNZA2GA	90
Serum albumin	RNJALBM	89
Fructose-1,6-bisphosphate 1-phosphohydrolase	RNFBP	88
Soares mouse melanoma (EST) (S[c])	AA124706	88
Soares mouse 3NbMS (EST) (AS[c])	AA154039	88

Table 2 (*Continued*)

FASTA-EMBL gene identification (rat unless otherwise stated)	Accession No.	Sequence homology[a] (%)
17-β-hydroxsteroid dehydrogenase	RN17BHDT2	87
Soares mouse p3NMF19.5 (EST)	AA038051	87
Peroxisomal enoyl-CoA:hydratase -3-hydroxyacyl CoA bifunctional enzyme	RNPECOA	85
Integral membrane protein, TAPA-1 (CD81) (mouse)	S45012	81
Soares mouse lymph node (EST)	MMAA88445	81
H. sapiens (clone zap128) mRNA	L40401	76
Lysophospholipase homologue (human)	HSU67963	76
Soares mouse lymph node (EST)	AA217044	74

[a] Refers to the nucleotide sequence homology between the cloned band isolated from the differential display and the corresponding gene derived from the EMBL gene sequence bank.

[b] EST is 'expressed sequence tag' — a gene of as yet unknown identity and function.

[c] Where sequence homologies were equal in both directions of the isolated band, both the sense (S) and antisense (A) identities are given.

Table 3
Identification of genes that were downregulated in male rat liver following 3-day treatment with Wy-14,643

FAST-EMBL gene identification (rat unless otherwise stated)	Accession No.	Sequence homology[a] (%)
NCI_CGAP_Li1 (*H. sapiens*) (EST[b])(S[c])	AA484528	99
NCI_CGAP_Pr1 (*H. sapiens*) (EST)(AS[c])	AA469320	99
UDP-glucuronosyltransferase (UGT2B12)	RN06273	98
Complement component c3	RNC3	96
Soares mouse placenta (S)	AA023305	96
Ape (chimpanzee) 28S rRNA (AS)	PTRGMC	96
Rat CYP2C11	RNCYPM1	95
Ribosomal protein S5	RNRPS5	94
Transthyretin	RNTTHY	94
Contrapsin-like protease inhibitor	RNCCP23	89
Prostaglandin F2a (S)	RN26663	84
β-2-microglobulin (AS)	RNB2MR	84
Apolipoprotein C-III	RNAPOA02	82
Parathymosin-alpha (zinc2+-binding protein)	RN11ZNBP	75

[a] Refers to the nucleotide sequence homology between the cloned band isolated from the differential display and the corresponding gene derived from the EMBL gene sequence bank.

[b] EST is 'expressed sequence tag' — a gene of as yet unknown identity and function.

[c] Where sequence homologies were equal in both directions, both the sense (S) and antisense (A) identities are given.

In all cases, both the forward and reverse sequence of the target clones were analysed and the gene having the highest statistical homology noted.

3.3. RT-PCR analysis of selected clones

The results of a typical RT-PCR semi-quantitation experiment for transferrin in the rat is given in Fig. 4 and the results for a total of 12 selected genes in both the rat and guinea pig are shown in Table 6.

4. Discussion

It is now apparent that all cancers arise from accumulated genetic changes within the cell. Although documenting and explaining these changes presents a formidable obstacle to understanding the different mechanisms of carcinogenesis, the experimental methodology is now available to begin attempting this difficult challenge. In order to begin the elucidation of the molecular mechanisms involved in non-genotoxic hepatocarcino-

ıesis, we have used SSH to identify a number of ıes that are upregulated or downregulated in ıle rat and guinea pig livers following short m exposure to the PP, Wy-14,643. We have ːd the rat model to represent a species suscepti- : to the non-genotoxic carcinogenic effect of 's and the guinea pig as a resistant species rton et al., 1984; Rodricks and Turnbull, 1987; Lake et al., 1989; Makowska et al., 1992; Lake et al., 1993).

Gurskaya et al. (1996), who originally developed the SSH technique, cloned the products of the secondary PCR reaction and screened a small number of randomly selected colonies for differentially expressed clones using northern hybridisation. However, we decided against this approach

ɔle 4
ntification of genes that were upregulated in male guinea pig liver following 3-day treatment with WY-14,643

STA-EMBL gene identification (guinea pig unless otherwise stated)	Accession No.	Sequence homology[a] (%)
·boxylesterase	AB010634	97
nplement C3 protein (GPC3)	M34054	97
osolic aldehyde dehydrogenase (sheep)	U12761	92
alase (human)	X04076	89
.ochondrial aspartate aminotransferase (pig)	M11732	89
ngation factor-1-alpha (rabbit)	X62245	88
I_CGAP_Br2 *H. sapiens* cDNA clone (EST) (Similar to chick mit. phosphoenolpyruate carboxykinase)	AA587436	87
·ha-1-antiproteinase S	M57270	83
'ormyltetrahydrofolate dehydrogenase (rat)	M59861	83
osomal protein L6 (rat)	X87107	83
res pregnant uterus Nb (EST) (mouse)	AA156847	83
ochondrial citrate transport protein (human)	L77567	80
oplasmic chaperonin hTRiC5 (human)	U17104	80
·ha-1-antiproteinase F	M57271	77
erogeneous nuclear ribonuclearprotein c1/c2 (human)	D28382	77
res parathyroid tumour (EST) (similar to human serum albumin precursor)	AA860651	76
ıtagene mouse kidney (EST)	AA107327	75
res parathyroid tumour NbHPA human cDNA (EST)	AA860653	74
res mouse mammary gland (EST)	AA619297	74
NA clone 15 004 (EST) (human)	H01826	74
res senescent fibroblasts (EST) (mouse)	W52190	74
proalbumin (human)	E04315	72
NA clone 73 169 (EST) (human)	T56624	72
ımin D-binding protein (human)	L10641	71
›H gene (exon 8) (human)	Y11498	71
⁽L flow sorted chromosome	B05457	71
res foetal liver spleen (EST) (mouse)	AA009524	71
res foetal heart NbMH19W (EST) (mouse)	AA009421	69
res foetal heart NbHH19W *H. sapiens* cDNA clone (EST)	W94377	67
nylalanine hydroxylase (human)	U49897	67
line-5-carboxylate dehydrogenase (human)	U24266	66
tathione-*S*-transferase homologue (human)	U90313	65
I_CGAP_GCBI (EST) (human)	AA769294	65
ıective protein (human)	M22960	64
ıe 27 375 (EST) (human)	N37046	62
tagene colon (# 937 204) *H. sapiens* cDNA clone (EST)	AA149777	62

Refers to the nucleotide sequence homology between the cloned band isolated from the differential display and the correspondgene derived from the EMBL gene sequence bank.

Table 5
Identification of genes that were downregulated in male guinea pig liver following 3-day treatment with WY-14,643

FASTA-EMBL gene identification (guinea pig unless otherwise stated)	Accession No.	Sequence homology[a] (%)
Complement C3 protein	M34054	97
Murinoglobulin	D84339	95
Alpha-1-antiproteinase F	M57271	88
Elongation factor-alpha-1 (rabbit)	X62245	89
Coupling protein G (human)	X04409	88
NCI_CGAP_Ov1 (EST[b]) (human)	AA586309	87
Lecithin:cholesterol acetyl transferase (rabbit)	D13668	85
Aldolase B (human)	X00270	84
Anti-thrombin III (human)	E00116	80
Phenylalanine hydroxylase (human)	K03020	80
Inter-α-trypsin inhibitor (human)	D38595	79
Normalised rat muscle (EST) (S[c])	AA849753	78
Normalised rat ovary (EST) (AS[c])	AA801059	78
Complement factor Ba fragment (human)	X00284	77
Dihydrodiol dehydrogenase (human)	U05598	76
Spot14 gene (thyroid-inducible hepatic protein)(human)	Y08409	75
BAC clone 174p12 (human)	AC004236	75
Mitochondrial aldehyde dehydrogenase (human)	X05409	74
Preproalbumin (human)	E04315	74
NCI_CGAP_Pr9 (EST) (human) (S)	AA533142	74
Normalised rat placenta (EST) (AS)	AA851197	74
Heparin sulfate proteoglycan (human)	J04621	73
cDNA clone 33 992 (EST) (human)	R24330	73
Retinol dehydrogenase (rat)	U33501	71
TAPA-1 integral membrane protein (CD81) (mouse)	S45012	71
Complement component c5s	M35525	70
Apolipoprotein B (pig)	L11235	69
cDNA clone 143 918 (EST) (human)	R76742	68
α-fibrinogen (human)	K02569	68
Soares foetal liver spleen 1NF (mouse)	W03726	68
Barstead bowel (EST) (mouse)	AA232049	67
UDP glucuronosyl transferase (cat)	AF0309137	66
Myeloid leukaemia cell differentiation protein (MCL-1) (human) (S)	L08246	65
STS SHGC-34 987 (human) (AS)	G27984	65
Soares mouse 3NME125	AA222798	64
Stratagene mouse embryonic (EST) (S)	AA199420	64
Rad 52 (mouse)	AF004854	63

[a] Refers to the nucleotide sequence homology between the cloned band isolated from the differential display and the corresponding gene derived from the EMBL gene sequence bank.
[b] EST is 'expressed sequence tag' — a gene of as yet unknown identity and function
[c] Where sequence homologies were equal in both directions, boththe sense (S) and antisense (A) identities are given.

for several reasons: (1) the kinetics of ligation and transformation favour the isolation of smaller PCR products, thereby producing a misrepresentation of larger gene products; (2) northern blot analysis is notoriously insensitive and is unlikely to confirm expression of rare transcripts; (3) there is no measurable end point to the screening of clones produced in this way other than to analyse every transformed colony. We used instead an alternative approach; after running out the differ-

ntial display on a high-resolution agarose gel Fig. 1) and overstaining with SYBR Green I to nhance visualisation, the composite bands were ndividually extracted, reamplified and cloned. Iowever, it has been well documented that single ands from differential displays often contain a eterogeneous mixture of different products Mathieu-Daude et al., 1996; Smith et al., 1997). This is because polyacrylamide gels cannot dis-riminate between DNA sequences that differ in ize by less than about 0.2% (Sambrook et al., 989). High-resolution agarose gels such as those ısed in this work are even less sensitive, normally nly discriminating products that differ in size by t least 1.5%. The use of the HA-red screening tep enables resolution of identical or nearly iden-ical sequences based on their AT content (Wawer t al., 1995) and is sensitive down to < 1% differ-nce. Furthermore, it is rapid, technically simple nd does not require the use of radiolabels. Geisinger et al. (1997) originally demonstrated the ısefulness of using HA-red to identify different roducts cloned from the same band of an RNA lifferential display experiment by simultaneously unning them in normal agarose (to discriminate y size) and in normal agarose containing HA-red to discriminate by AT content). We have found hat this approach is equally useful for identifying ifferent gene species cloned from the same band f our SSH display.

Diatchenko et al. (1996) reported that SSH is ighly efficient at producing differentially ex-ressed gene species. However, we also included a econd screening step to further confirm that the lones isolated from the differential display were ndeed differentially expressed. Duplicate dotblots f the candidate clones were blotted with the isplay from which they were originally isolated nd with the 'reverse subtraction' display. To nake the reverse-subtracted probe, the subtractive ybridisation step of the procedure was carried ut using the original tester cDNA as a driver, nd the original driver cDNA as a tester. In this vay, clones that are false positives can be iden-ified through their presence in both blots. Such alse positives most commonly arise through hav-ng a very high abundance in the initial sample or nusual hybridisation properties (Li et al., 1994).

Although the SSH method itself has been shown to be efficient, and despite the screening step that we included, there is an important caveat to bear in mind — namely that it is important that all clones be considered only as 'candidates' until the actual abundance of their mRNA is quantitated in treated and control samples. Towards this end, we examined the expression of a limited number of clones using semi-quantitative RT-PCR. Albumin was used as the reference gene as we have previously found that the expression of this gene does not appear to change with the treatment regime that we used (Fig. 4, and data not shown). There are a number of interesting points to note from our results. The first is the presence of genes that serve as appropriate positive controls in the upregulated and downregulated series. For example, in the rat it can be seen that CYP4AI expression increases 14-fold following treatment. Although CYP4AI mRNA expression levels following WY-14,643 treatment have not been previously reported in this model, the figure compares favourably with that recorded by Bell et al. (1991), who used RNAse-protection to quantitate CYP4A1 in rat liver following treatment with methylclofenapate, another PP. In addition, we also confirmed that the peroxisomal enoyl-CoA:hydratase-3-hydroxyacyl-CoA bifunctional enzyme is also upregulated 9-fold, in agreement with the findings of Chen and Crane (1992).

A number of genes were downregulated following Wy-14,643 exposure, including CYP2C11 expression. Corton et al. (1997) reported similar findings and suggested that this may in part explain why male rats exposed to Wy-14,643 and some other PPs have high serum estradiol levels, as estradiol is a substrate for CYP2C11. We have also shown that the expression of contrapsin-like protease inhibitor (CLPI) was downregulated by Wy-14,643. This has not previously been reported, and we suggest that it may be linked to a requirement for increased availability of amino acids to accommodate the hepatomegaly induced by treatment. Although little is known of the function of parathymosin-α, (zinc^{2+}-binding protein) it has been shown to interact with the globular domain of histone H1, suggesting a role in histone function (Kondili et al., 1996). In contrast to the

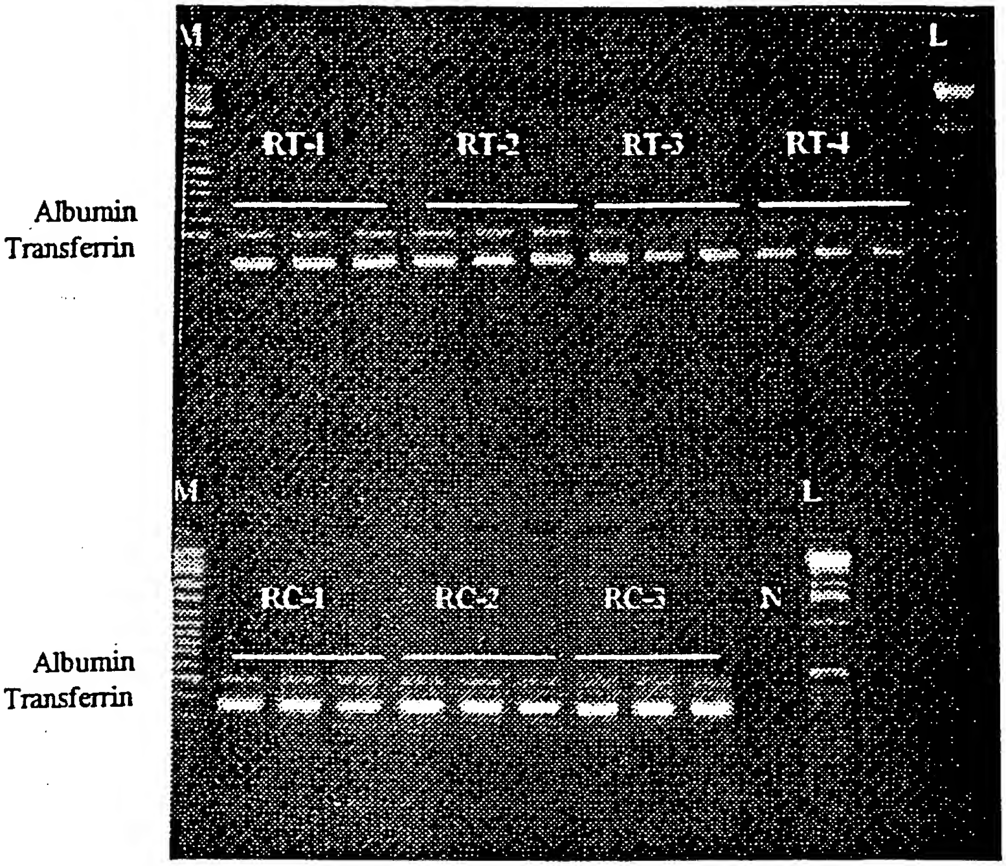


Fig. 4. Semi-quantitative RT-PCR experiment showing relative decrease in expression of transferrin in treated rat liver (RT-1 to RT-4) compared to controls (RC-1 to RC-3). An equal amount of mRNA was used in each reaction (10 ng), and each sample was quantitated in triplicate to reduce the effects of inter-tube variation. N is negative control (no mRNA). Lane M is a 100 bp ladder and lane L is a 1 Kb DNA ladder.

downregulation observed in this work, other studies have shown that parathymosin-α expression is elevated in breast cancer (Tsitsilonis et al., 1993, 1998), with the implication that parathymosin-α may somehow be involved in regulating cell proliferation by more than one mechanism. Transferrin has previously been shown to be downregulated in rat liver by hypolipidemic PPs (Hertz et al., 1996). It is therefore interesting to note that we isolated a clone identified as transferrin from the upregulated display profile. Since we confirmed by RT-PCR that transferrin is in fact downregulated in the rat (Fig. 4), we conclude that transferrin was either a false positive or was incorrectly identified. It could also be that we have isolated a close relative, splice variant or isoform of transferrin, which demonstrates a different expression profile under these experimental conditions. Further investigations are therefore required to determine which of these possibilities are correct.

One of our most intriguing observations was that one gene, CD81, appeared to be upregulated in rat liver but downregulated in guinea pig liver following Wy-14,643 exposure. CD81 is a widely expressed cell surface protein that is involved in a large number of cellular functions, including adhesion, activation, proliferation and differentiation (reviewed by Levy et al., 1998). Since all of these functions are altered to some extent in carcinogenesis, it is perhaps an important observation that CD81 expression is differentially regulated in a resistant and sensitive species exposed to a non-genotoxic carcinogen.

Albumin and ribosomal genes appear common to all differential displays and are thus undesirable false positives. However, due to their high expression in the liver, they are difficult to re-

ıove. We also noted a number of gene species, articularly in the guinea pig, which were comıon to both upregulated and downregulated rofiles. Again, the most likely reason for these aving arisen is their high abundance.

A relatively large number of upregulated and ownregulated genes were isolated from guinea ig liver following Wy-14,643 exposure. However, ıe guinea pig genome has been relatively poorly haracterised and so many of the clones were lentified as resembling genes or ESTs from other pecies. Without full-length sequence data it is ifficult to ascertain the accuracy of the assigned lentities and this must be borne in mind when tilising data such as this, for example, in designıg effective primers for RT-PCR studies. Alıough the actual isolated clone sequences can be sed to do this, their relatively small size often estricts the ability to design effective primers. In ddition, as we observed with transferrin, using a ublished full-length sequence may help to idenfy false positives.

By comparing the expression profiles of genes showing altered expression in a PP-sensitive species (rat) with a PP-resistant species (guinea pig), it was our aim to identify genes that are mechanistically relevant to the non-genotoxic hepatocarcinogenic action of Wy-14,643. However, few of the genes that we have isolated were common to both the rat and the guinea pig. This suggests either that the molecular mechanisms of response in these two species are so different that few genes are commonly regulated in response to Wy-14,643 exposure, or that we have recovered only a small proportion of those genes that have altered expression. The latter seems the more likely scenario since it is perceived that one of the main problems of subtractive hybridisation and other differential expression technologies is the inability to consistently isolate rare gene transcripts (Bertioli et al., 1995). This is potentially problematic in that weakly expressed genes may play an important role in regulating key cellular processes, and that the majority of mRNA species are classified as

able 6
emi-quantitative RT-PCR analysis of selected gene species in the rat and guinea pig[a]

ranscript	Putative change of expression following treatment according to dotblot		Change according to RT-PCR quantitation	
	Rat	Guinea pig	Rat	Guinea pig
lbumin	N/A	N/A	No change	No change
ifunctional enzyme	Up	N/A	Upregulated* (9 ×)	N/O
YP2C11	Down	N/A	Downregulated* (Abolished)	N/D
YP4A1	Up	N/A	Upregulated* (14 ×)	N/D
atalase	N/A	Up	No change	N/O
D81 (TAPA-1)	Up	Down	N/O	Upregulated**(1.4 ×)
ontrapsin-like protease inhibitor	Down	N/A	Downregulated** (0.5 ×)	N/D
arathymosin-α (zinc^{2+} binding protein)	Down	N/A	Downregulated** (0.6 ×)	N/D
ransferrin	Up	N/A	Downregulated* (0.5 ×)	No change
DP-Glucuronosyl transferase	Down	N/A	Downregulated** (0.2 ×)	N/O
ownUnknown-1	Down	N/A	No change ($P = 0.06$)	N/D
ı-α2-glycoprotein	Up	N/A	No change	N/O

[a] N/A, not applicable; N/O, not optimised; N/D, not done.
* $P < 0.0005$;
** $P < 0.05$.

'rare' in abundance (Bertioli et al., 1995). However, in their original paper describing the SSH technique, Gurskaya et al. (1996) demonstrated that SSH can enrich rare molecules between 1000- and 5000-fold in a single round of hybridisation. Unfortunately, due to high background smearing in our initial experiments (which hindered identification of single bands), we were compelled to reduce the primary hybridisation time to only 4 h — a step that theoretically is likely to reduce the number of rare sequences (CLONTECHniques, 1996). Furthermore, it has been claimed by the manufacturers that, whilst this technique can identify changes as small as 1.5-fold between the driver and tester populations, it is best suited to the isolation of genes that show a greater than 5-fold increase (CLONTECHniques, 1996). In addition, where tester and driver contain genes with large and small differences in abundance, the SSH method will be biased towards identifying those genes with the large differences (CLONTECHniques, 1996). Thus, it is most probable that we have not isolated all of the more rarely expressed transcripts and those demonstrating small changes in expression.

One problem that remains is identifying the function of genes isolated in SSH experiments as described herein, some of which may be crucial to the process of carcinogenesis, and are, to date, unidentified. However, we have provided evidence herein that SSH can be used to begin the process of characterising the extent and importance of altered gene expression in response to a chemical stimulus. The developments of this approach should include characterisation of temporal and dose responses, and functional analysis studies including knockout mice. In combination, such studies should make a significant contribution to our understanding of the molecular mechanisms of action and physiological relevance of gene regulation in non-genotoxic hepatocarcinogenesis. It should then be possible to ascertain whether differentially expressed genes are causally or casually related to the chemical-induced toxicity, and therefore a substantial mechanistic advance.

It is clear that there are also broader applications for this experimental approach that go beyond understanding the molecular mechanisms of peroxisome-proliferator induced non-genotoxic hepatocarcinogenesis in rodents. The potential medical and therapeutic benefits of elucidating the molecular changes that occur in any given cell in progressing from the normal to the carcinogenic (or other diseased, abnormal or developmental) state are very substantial. Notwithstanding the lack of complete functional identification of altered gene expression, such gene profiling studies described herein essentially provides a 'fingerprint' of each stage of carcinogenesis, and should help in the elucidation of specific and sensitive biomarkers for different types of cancer. Amongst other benefits, such fingerprints and biomarkers could help uncover differences in histologically identical cancers, and provide diagnostic tests for the earliest stages of neoplasia. In addition, the genes identified by this approach may be incorporated into gene-chip DNA-arrays, thus providing a standard genetic fingerprint for a particular toxin treatment in a particular species. Interrogation of these gene arrays for an unknown compound that has a similar pattern to the known reference chemical would then provide evidence that the unknown may have a toxicity profile similar to the 'standard' fingerprint, thereby serving as a mechanistically relevant platform for further detailed investigations.

Acknowledgements

This work was funded by Rhone-Poulenc Agrochemicals, Nice, France.

References

Anderson, N.L., Esquer-Blasco, R., Richardson, F., Foxworthy, P., Eacho, P., 1996. The effects of peroxisome proliferators on protein abundances in mouse liver. Toxicol. Appl. Pharmacol. 137, 75–89.

Ashby, J., 1992. Prediction of non-genotoxic carcinogenesis. Toxicol. Lett. 64–65, 605–612.

Bell, D.R., Bars, R.G., Gibson, G.G., Elcombe, C.R., 1991. Localisation and differential induction of cytochrome P450IVA and acyl coA oxidase in rat liver. Biochem. J. 275, 247–252.

Bertioli, D.J., Schlichter, U.H.A., Adams, M.J., Burrows, P.R., Steinbiss, H.-H., Antoniw, J.F., 1995. An analysis of

differential display shows a strong bias towards high copy number mRNAs. Nucleic Acid Res. 23 (21), 4520–4523.

Cattley, R.C., Kato, M., Popp, J.A., Teets, V.J., Voss, K.S., 1994. Initiator-specific promotion of hepatocarcinogenesis by Wy-14,643 and clofibrate. Carcinogenesis 15 (8), 1763–1766.

Chen, N., Crane, D.I., 1992. Induction of the major integral membrane protein of mouse liver peroxisomes by peroxisome proliferators. Biochem J. 283, 605–610.

CLONTECHniques, 1996. Technical Tips: Clontech PCR-Select cDNA Subtraction, October 25, application notes.

CLONTECHniques, 1997. PCR-Select Differential Screening Kit — The Next Step After Clontech PCR-Select cDNA Subtraction. XII(2), 18–19, application notes.

Corton, J.C., Bocos, C., Moreno, E.S., Merrit, A., Cattley, R.C., Gustaffson, J.A., 1997. Peroxisome proliferators alter the expression of estrogen-metabolising enzymes. Biochimie 79, 151–162.

Davis, M., Cohen, D.I., Nielson, E.A., Steinmetz, M., Paul, W.E., Hood, L., 1984. Cell-type-specific cDNA probes and the murine I region: the localisation and orientation of Ad/a. Proc. Natl. Acad. Sci USA 81, 2194–2198.

Diatchenko, L., Lau, Y.-F.C., Campbell, A.P., Chenchik, A., Moqadam, F., Huang, B., Lukyanov, K., Gurskaya, N., Sverdlov, E.D., Siebert, P.D., 1996. Suppression subtractive hybridisation: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. Proc. Natl. Acad. Sci. USA 93, 6025–6030.

Duguid, J., Dinauer, M., 1990. Library subtraction of in vitro cDNA libraries to identify differentially expressed genes in scrapie infection. Nucleic Acid Res. 18 (9), 2789–2792.

Geisinger, A., Rodriguez, R., Romero, V., Wettstein, R., 1997. A simple method for screening cDNAs arising from the cloning of RNA differential display bands. Elsevier trends journals technical tips online, http://tto.trends.com, document number T01110

Guimaraes, M.J., Lee, F., Zlotnik, A., McClanahan, T., 1995. Differential display by PCR: novel findings and applications. Nucleic Acid Res. 23 (10), 1832–1833.

Gurskaya, N.G., Diatchenko, L., Chenchik, P.D., Siebert, P.D., Khaspekov, G.L., Lukyanov, K.A., Vagner, L.L., Ermolaeva, O.D., Lukyanov, S.A., Sverdlov, E.D., 1996. Equalising cDNA subtraction based on selective suppression of polymerase chain reaction: cloning of Jurkat cell transcripts induced by phytohemaglutinin and phorbol 12-myrystate 13-acetate. Anal. Biochem. 240, 90–97.

Hayashi, F., Tamura, H., Yamada, J., Kasai, H., Suga, T., 1994. Characteristics of the hepatocarcinogenesis caused by dehydroepiandrosterone, a peroxisome proliferator, in male F-344 rats. Carcinogenesis 15 (190), 2215–2219.

Hedrick, S.M., Cohen, D.I., Nielsen, E.A., Davis, M.M., 1984. Isolation of cDNA clones encoding T cell-specific membrane-associated proteins. Nature 308 (8), 149–153.

Hertz, R., Seckbach, M., Zakin, M.M., Bar-Tana, J., 1996. Transcriptional suppression of the transferin gene by hypolipidemic peroxisome proliferators. J. Biol. Chem. 271 (1), 218–224.

Hubank, M., Schatz, D.G., 1994. Identifying differences in mRNA expression by representational difference analysis. Nucleic Acid Res. 22 (25), 5640–5648.

Human and Experimental Toxicology, 1994. Hum. Exp. Toxicol. 13 (Suppl. 2) (entire issue).

Kolls, J., Dsininger, P., Cohen, C., Larson, J., 1993. cDNA equalisation for reverse transcription-polymerase chain reaction quantitation. Anal. Biochem 208, 264–269.

Kondili, K., Tsolas, O., Papamarcaki, T., 1996. Selective interaction between parathymosin and histone H1. Eur. J. Biochem. 242 (1), 67–74.

Lake, B.G., Evans, J.G., Gray, T.J.B., Korosi, S.A., North, C.J., 1989. Comparative studies of nafenopin-induced hepatic peroxisome proliferation in the rat, Syrian hamster, guiea pig and marmoset. Toxicol. Appl. Pharmacol. 99, 148–160.

Lake, B.G., Evans, J.G., Cunninghame, M.E., Price, R.J., 1993. Comparison of the hepatic effects of Wy-14,643 on peroxisome proliferation and cell replication in the rat and Syrian hamster. Environ. Health Perspect. 101 (S5), 241–248.

Levy, S., Todd, S.C., Maecker, H.T., 1998. CD81 (TAPA-1): a molecule involved in signal transduction and cell adhesion in the immune system. Annu. Rev. Immunol. 16, 89–109.

Li, W.B., Gruber, C.E., Lin, J.J., D'Alessio, J.M., Jessee, J.A., 1994. The isolation of differentially expressed genes in fibroblast growth factor stimulated BC3H1 cells by subtractive hybridization. BioTechniques 16, 722–729.

Liang, P., Pardee, A.B., 1992. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. Science 257 (5072), 967–971.

Liang, P., Averboukh, L., Keyomarsi, K., Sager, R., Pardee, A.B., 1992. Differential display and cloning of messenger RNAs from human breast cancer versus mammary epithelium. Cancer Res. 52, 6966–6968.

Lipman, D.J., Pearson, W.R., 1985. Rapid and sensitive protein similarity searches. Science 227, 1435–1441.

Makowska, J.M., Gibson, G.G., Bonner, F.W., 1992. Species differences in ciprofibrate induction of hepaic cytochrome P450IVA1 and peroxisome proliferation. J. Biochem. Toxicol. 7, 183–191.

Marsman, D.S., Cattley, R.C., Conway, J.G., Popp, J.A., 1988. Relationship of hepatic peroxisome proliferation and replicative DNA synthesis to the hepatocarcinogenicity of the peroxisome proliferators di-(2-ethylhexyl)phthalate and [4-chloro-6-(2,3-xylidino)-2-pyrimidinylthio]acetic acid (Wy-14,643) in rats. Cancer Res. 48, 6739–6744.

Mathieu-Daude, F., Cheng, R., Welsh, J., McClelland, M., 1996. Screening of differentially amplified cDNA products from RNA arbitrarily primed PCR fingerprints using single strand conformation polymorphism (SSCP) gels. Nucleic Acid Res. 24 (8), 1504–1507.

Orton, T.C., Adam, H.K., Bentley, M., Holloway, B., Tucker, M.J., 1984. Clobuzarit: species differences in the morphological and biochemical response of the liver following chronic administration. Toxicol. Appl. Pharmacol. 73, 138–151.

Parodi, S., 1992. Non-genotoxic factors in the carcinogenic process: problems of detection and hazard evaluation. Toxicol. Lett. 64–65, 621–630.

Pearson, W.R., Lipman, D.J., 1988. Imported tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA 85, 2444–2448.

Rockett, J.C., Esdaile, D.J., Gibson, G.G., 1997. Molecular profiling of non-genotoxic hepatocarcinogenesis using differential display reverse transcription-polymerase chain reaction (ddRT-PCR). Eur. J. Drug. Metab. Pharmacokinet 22 (4), 329–333.

Rodricks, J.V., Turnbull, D., 1987. Inter-species differences in peroxisomes and peroxisome proliferation. Toxicol. Ind. Health 3, 197–212.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. In: Ford, N., Nolan, C., Ferguson, M. (Eds.), Molecular Cloning — A Laboratory Manual, second ed. Cold Spring Harbor Laboratory Press, New York.

Sargent, T., Dawid, I., 1983. Differential gene expression in the gastrula of *Xenopus laevis*. Science 222, 135–139.

Smith, N.R., Li, A., Aldersley, M., High, A.s., Markham, A.F., Robinson, P.A., 1997. Rapid determination of the complexity of cDNA bands extracted from DDRT-PCR polyacrylamide gels. Nucleic Acid Res. 25 (17), 3552–3554.

Tsitsilonis, O.E., Stiakakis, J., Koutselinis, A., Gogas, J., Markopoulos, C., Yialouris, P., Bekris, S., Panoussopoulos, D., Kiortsis, V., Voelter, W., Haritos, A.A., 1993. Expression of alpha-thymosins in human tissues in normal and abnormal growth. Proc. Natl. Acad. Sci. USA 90 (20), 9504–9507.

Tsitsilonis, O.E., Bekris, E., Voutsas, I.F., Baxevanis, C.N., Markopoulos, C., Papadopoulou, S.A., Kontzoglou, K., Stoeva, S., Gogas, J., Voelter, W., Papamichail, M., 1998. The prognostic value of alpha-thymosins in breast cancer. Anticancer Res. 18 (3A), 1501–1508.

Wan, J.S., Sharp, S.J., Poirier, G.M.-C., Wagaman, P.C., Chambers, J., Pyati, J., Hom, Y.-L., Galindo, J.E., Huvar, A., Peterson, P.A., Jackson, M.R., Erlander, M.G., 1996. Cloning differentially expressed mRNAs. Nat. Biotechnol. 14, 1685–1691.

Wawer, C., Ruggeberg, H., Meyer, G., Muyzer, G., 1995. A simple and rapid electrophoresis method to detect sequence variation in PCR-amplified DNA fragments. Nucleic Acid Res. 23 (23), 4928–4929.

TOXICOLOGY

An international journal concerned with the effects of chemicals on living systems and immunotoxicology

ELSEVIER

Special Issue

Festschrift dedicated to Professor Dr. K.J. Netter

Proc. Natl. Acad. Sci. USA
Vol. 94, pp. 13057–13062, November 1997
Genetics

Yeast microarrays for genome wide parallel genetic and gene expression analysis

DEVAL A. LASHKARI*†, JOSEPH L. DERISI‡, JOHN H. MCCUSKER§, ALLEN F. NAMATH‡, CRISTL GENTILE§, SEUNG Y. HWANG‡, PATRICK O. BROWN‡, AND RONALD W. DAVIS*‡

Departments of *Genetics and ‡Biochemistry, Stanford University, Stanford, CA 94305; and §Department of Microbiology, Duke University, Durham, NC 27710

Contributed by Ronald W. Davis, September 2, 1997

ABSTRACT **We have developed high-density DNA microarrays of yeast ORFs. These microarrays can monitor hybridization to ORFs for applications such as quantitative differential gene expression analysis and screening for sequence polymorphisms. Automated scripts retrieved sequence information from public databases to locate predicted ORFs and select appropriate primers for amplification. The primers were used to amplify yeast ORFs in 96-well plates, and the resulting products were arrayed using an automated micro arraying device. Arrays containing up to 2,479 yeast ORFs were printed on a single slide. The hybridization of fluorescently labeled samples to the array were detected and quantitated with a laser confocal scanning microscope. Applications of the microarrays are shown for genetic and gene expression analysis at the whole genome level.**

The genome sequencing projects have generated and will continue to generate enormous amounts of sequence data. The genomes of *Saccharomyces cerevisiae, Haemophilus influenzae* (1), *Mycoplasma genitalium* (2), and *Methanococcus jannischii* (3) have been completely sequenced. Other model organisms have had substantial portions of their genomes sequenced as well including the nematode *Caenorhabditis elegans* (4) and the small flowering plant *Arabidopsis thaliana* (5). Given this ever-increasing amount of sequence information, new strategies are necessary to efficiently pursue the next phase of the genome projects—the elucidation of gene expression patterns and gene product function on a whole genome scale.

One important use of genome sequence data is to attempt to identify the functions of predicted ORFs within the genome. Many of the ORFs identified in the yeast genome sequence were not identified in decades of genetic studies and have no significant homology to previously identified sequences in the database. In addition, even in cases where ORFs have significant homology to sequences in the database, or have known sequence motifs (e.g., protein kinase), this is not sufficient to determine the actual biological role of the gene product. Experimental analysis must be performed to thoroughly understand the biological function of a given ORF's product. Model organisms, such as *S. cerevisiae*, will be extremely important in improving our understanding of other more complex and less manipulable organisms.

To examine in detail the functional role of individual ORFs and relationships between genes at the expression level, this work describes the use of genome sequence information to study large numbers of genes efficiently and systematically. The procedure was as follows. (*i*) Software scripts scanned annotated sequence information from public databases for predicted ORFs. (*ii*) The start and stop position of each identified ORF was extracted automatically, along with the sequence data of the ORF and 200 bases flanking either side. (*iii*) These data were used to automatically select PCR primers that would amplify the ORF. (*iv*) The primer sequences were automatically input into the automated multiplex oligonucleotide synthesizer (6). (*v*) The oligonucleotides were synthesized in 96-well format, and (*vi*) used in 96-well format to amplify the desired ORFs from a genomic DNA template. (*vii*) The products were arrayed using a high-density DNA arrayer (7–10). The gene arrays can be used for hybridization with a variety of labeled products such as cDNA for gene expression analysis or genomic DNA for strain comparisons, and genomic mismatch scanning purified DNA for genotyping (11).

METHODS

Script Design. All scripts were written in UNIX Tool Command Language. Annotated sequence information from GenBank was extracted into one file containing the complete nucleotide sequence of a single chromosome. A second file contained the assigned ORF name followed by the start and stop positions of that ORF. The actual sequence contained within the specified range, along with 200 bases of sequence flanking both sides, was extracted and input into the primer selection program PRIMER 0.5 (Whitehead Institute, Boston). Primers were designed so as to allow amplification of entire ORFs. The selected primer sequences were read by the 96-well automated multiplex oligonucleotide synthesizer instrument for primer synthesis. The forward and reverse primers were synthesized in two separate 96-well plates in corresponding wells. All primers were synthesized on a 20-nmol scale.

ORF Amplification and Purification. Genomic DNA was isolated as described (12) and used as template for the amplification reactions. Each PCR was done in a total volume of 100 μl. A total of 0.2 μM each of forward and reverse primers were aliquoted into a 96-well PCR plate (Robbins Scientific, Sunnyvale, CA); a master mix containing 0.24 mM each dNTP, 10 mM Tris (pH 8.5), 50 mM $MgCl_2$, 2.5 units *Taq* polymerase, and 10 ng of template was added to the primers, and the entire mix was thermal cycled for 30 cycles as follows: 15 min at 94°C, 15 min at 54°C, and 30 min at 72°C. Products were ethanol precipitated in polystyrene v-bottom 96-well plates (Costar). All samples were dried and stored at −20°C.

Arraying Procedure and Processing. Microarrays were made as described (8).

A custom built arraying robot was used to print batches of 48 slides. The robot utilizes four printing tips which simultaneously pick up ≈1 μl of solution from 96-well microtiter plates. After printing, the microarrays were rehydrated for 30 sec in a humid chamber and then snap dried for 2 sec on a hot plate (100°C). The DNA was then UV crosslinked to the surface by subjecting the slides to 60 millijoules of energy. The rest of the poly-L-lysine surface was blocked by a 15-min incubation in a solution of 70 mM succinic anhydride dissolved in a solution consisting of 315 ml of 1-methyl-2-pyrrolidinone (Aldrich) and 35 ml of 1 M boric acid (pH 8.0). Directly after the blocking reaction, the bound DNA was denatured by a 2-min incubation in distilled water at ≈95°C.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: YEP, yeast extract/peptone.

†To whom reprint requests should be sent at the present address: Synteni, Inc., 6519 Dumbarton Circle, Fremont, CA 94555.

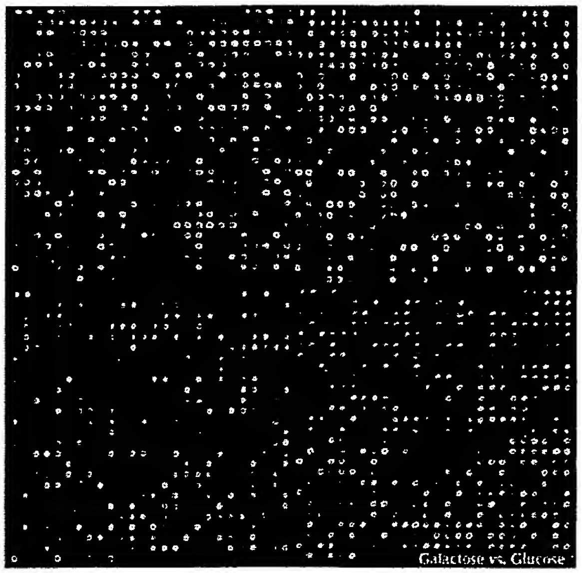


FIG. 1. Two-color fluorescent scan of a yeast microarray containing 2,479 elements (ORFs). The center-to-center distance between elements is 345 μm. A probe mixture consisting of cDNA from yeast extract/peptone (YEP) galactose (green pseudocolor) and YEP glucose (red pseudocolor) grown yeast cultures was hybridized to the array. Intensity per element corresponds to ORF expression, and pseudocolor per element corresponds to relative ORF expression between the two cultures.

The slides were then transferred into a bath of 100% ethanol at room temperature.

Probe Preparation: cDNA. Yeast cultures (100 ml) were grown to ≈1 OD_{A600} and total RNA was isolated as described (13). Up to 500 μg total RNA was used to isolate mRNA (Qiagen, Chatsworth, CA). Oligo(dT)20 (5 μg) was added and annealed to 2 μg of mRNA by heating the reaction to 70°C for 10 min and quick chilling on ice, plus 2 μl SuperScript II (200 units/μl) (Life Technologies, Gaithersburg, MD), 0.6 μl 50× dNTP mix (final concentrations were 500 μM dATP, dCTP, dGTP, and 200 μM dTTP), 6 μl 5× reaction buffer, and 60 μM Cy3-dUTP or Cy5-dUTP (Amersham). Reactions were carried out at 42°C for 2 h, after which the mRNA was degraded by the addition of 0.3 μl 5 M NaOH and 0.3 μl 100 mM EDTA and heating to 65°C for 10 min. The sample was then diluted to 500 μl with TE and concentrated using a Microcon-30 (Amicon) to 10 μl.

Probe Preparation: Genomic DNA. Fluorescent DNA was prepared from total genomic DNA as follows: 1 μg of random nonamer oligonucleotides was added to 2.5 μg of genomic DNA. This mixture was boiled for 2 min and then chilled on ice. A reaction mixture containing dNTPs (25 μM dATP, dCTP, dGTP, 10 μM dTTP, and 40 μM Cy3-dUTP or Cy5-dUTP) reaction buffer (New England Biolabs), and 20 units exonuclease free Klenow enzyme (United States Biochemical) was added, and the reaction was incubated at 37°C for 2 h. The sample was then diluted to 500 μl with TE and concentrated using a Microcon-30 (Amicon) to 10 μl.

Hybridization. Purified, labeled probe was resuspended in 11 μl of 3.5× SSC containing 10 μg *Escherichia coli* tRNA, and 0.3% SDS. The sample was then heated for 2 min in boiling water, cooled rapidly to room temperature, and applied to the array. The array was placed in a sealed, humidified, hybridization chamber. Hybridization was carried out for 10 h in a 62°C water bath, after which the arrays were washed immediately in 2× SSC/0.2% SDS. A second wash was performed in 0.1× SSC.

Analysis and Quantitation. Arrays were scanned on a scanning laser fluorescence microscope developed by Steve Smith with software written by Noam Ziv (Stanford University). A separate scan was done for each of the two fluorophores used. The images were then combined for analysis. A bounding box, fitted to the size of the DNA spots, was placed over each array element. The average fluorescent intensity was calculated by summing the intensities of each pixel present in a bounding box and then dividing by the total number of pixels. Local area background was calculated for each array element by determining the average fluorescent intensity at the edge of the bounding box. To normalize for fluorophore-specific variation, control spots containing yeast genomic DNA were applied to each quadrant during the arraying process. These elements were quantitated and the ratios of the signals were determined. These ratios were then used to normalize the photomultiplier sensitivity settings such that the ratios of the fluorescence of the genomic DNA spots were close to a value of 1.0. The average signal intensity at any given spot was regarded as significant if it was at least two standard deviations above background. Each experiment was conducted in duplicate, with the fluorophores representing each channel reversed. The ratios presented here are the average of the two experiments, except in the case in which the signal for the element in question was below the reliability threshold. The reliability threshold also determined the dynamic range of the experiment. For all of the experiments presented, the average dynamic range was ≈1 to 100. In the case where the fluorescence from a very bright spot saturates the detector, differential ratios will, in general, be underestimated. This can be compensated for by scanning at a lower overall sensitivity.

RESULTS

The accumulation of sequence information from model organisms presents an enormous opportunity and challenge to understand the biological function of many previously uncharacterized genes. To do this accurately and efficiently, a directed strategy was developed that enables the monitoring of multiple genes simultaneously. Microarraying technology provides a method by which DNA can be attached to a glass surface in a high-density format (8). In practice, it is possible to array over 6,000 elements in an area less than 1.8 cm^2. Given that the yeast genome consists of ≈6,100 ORFs, the entire set of yeast genes can be spotted onto a single glass slide.

With this capability and the availability of the entire sequence of the yeast genome, our strategy was to use a directed approach for generating the complete genome array. This procedure involved synthesizing a pair of oligonucleotide primers to amplify each ORF. The PCR product containing each gene of interest was arrayed onto glass and used, for example, as probe for monitoring gene expression levels by hybridizing to the array labeled cDNA generated from isolated mRNA of a culture grown under any experimental condition.

Primer Selection and Synthesis. The primer selection was fully automated using Tool Command Language scripts and PRIMER 0.5. (Whitehead). Primer pairs were automatically selected successfully for >99% of the ORFs tested. Primer sequences can thus be selected rapidly with minimal manual processing. A complete set of forward and reverse primers were selected initially for each ORF on chromosomes I, II, III, V, VI, VIII, IX, X, and XI. Primers for a representative set of ORFs (15% coverage) were chosen for the remaining chromosomes. With the release of the entire yeast genome sequence, the complete set of primers has now been selected.

Because each ORF requires a unique pair of synthetic primers, a total of approximately 12,200 oligonucleotides will be required to individually amplify each target. This costly component was addressed with the automated multiplex oligonucleotide synthesizer (6) which efficiently synthesizes primers in a 96-well format. Each primer, synthesized on a 20-nmol scale, provides enough material for 100 amplification reactions, whereas a given PCR product provides enough material to generate an element on

Table 1. Heat shock vs. control expression data

Ratio of gene expression				
Control	Heat	ORF	Gene	Description
	2.2	YLR142	PUT1	Proline oxidase
	2.0	YOL140	ARG8	Acetylornithine aminotransferase
2.3		YGL148	ARO2	Chorismate synthase
	36.0	YFL014	HSP12	Heat shock protein
	27.4	YBR072	HSP26	Heat shock protein
	6.7	YBR054	YRO2	Similarity to HSP30 heat shock protein Yro1p
	3.4	YCR021	HSP30	Heat shock protein
	2.6	YER103	SSA4	Heat shock protein
	2.5	YLR259	HSP60	Mitochondrial heat shock protein HSP60
	2.1	YBR169	SSE2	Heat shock protein of the HSP70 family
	1.7	YBL075	SSA3	Cytoplasmic heat shock protein
	1.4	YPL240	HSP82	Heat shock protein
	1.4	YDR258	HSP78	Mitochondrial heat shock protein of clpb family of ATP-dependent proteases
1.0		YNL007	SIS1	Heat shock protein
1.1		YEL030		70-kDa heat shock protein
1.9		YHR064		Heat shock protein
	1.3	YBL008	HIR1	Histone transcription regulator
2.6		YBL002	HTB2	Histone H2B.2
3.3		YBL003	HTA2	Histone H2A.2
3.3		YBR010	HHT1	Histone H3
3.9		YBR009	HHF1	Histone H4
	2.4	YDR343	HXT6	High-affinity hexose transporter
	2.1	YHR092	HXT4	Moderate- to low-affinity glucose transporter
3.6		YAR071	PHO11	Secreted acid phosphatase, 56 kDa isozyme
	2.3	YLR096	KIN2	Ser/Thr protein kinase
2.5		YER102	RPS8B	Ribosomal protein S8.e
2.6		YBR181	RPS101	Ribosomal protein S6.e
2.6		YCR031	CRY1	40S ribosomal protein S14.e
2.7		YLR441	RP10A	Ribosomal protein S3.a.e
2.8		YHR141	RPL41B	Ribosomal protein L36a.e
2.8		YBL072	RPS8A	Ribosomal protein S8.e
2.8		YHL015	URP2	Ribosomal protein
2.8		YBR191	URP1	Ribosomal protein L21.e
3.1		YLR340	RPLA0	Acidic Ribosomal protein L10.e
3.3		YGL123	SUP44	Ribosomal protein
	5.8	YLR194		Hypothetical protein

500–1,000 arrays. Thus, a single primer pair provides enough starting material for up to ≈50,000 arrays.

Primers were synthesized to amplify yeast ORFs. Primer synthesis had a failure rate of <1% in over 18 plates of synthesis as determined by standard trityl analysis (6). The success rate of the PCR amplifications using the primer pairs was 94% based on agarose gel analysis of each PCR. The purified PCR products were used to generate arrays. Two versions of the arrays were created for the experimental results presented here. The first array contained 2,287 elements and the second array batch contained 2,479 elements.

Genome Arrays. The amplified ORFs were arrayed onto glass at a spacing of 345 microns (Fig. 1). The high-density spacing of DNA samples allows the hybridization volumes to be minimized—volumes are a maximum of 10 μl. The labeled probe can thus be maintained at relatively high concentrations, making 1–2 μg of mRNA sufficient for analysis. This also obviates the need for a subsequent amplification step and thus avoids the risk of altering the relative ratios of different cDNA species in the sample.

Genetic Analysis: Genomic Comparison of Unrelated Strains. Microarrays allow efficient comparison of the genomes of different strains. Genomic DNA from Y55, an *S. cerevisiae* strain divergent from the reference strain S288c, was randomly labeled with Cy3-dUTP and hybridized simultaneously with the S288c DNA labeled with Cy5-dUTP. When a comparison between the hybridization of the DNA from the two strains was done, several elements gave relatively little or no signal above background from the Cy3 channel (data not shown). These include SGE1, ASP3A-D, YLR156, YLR159, YLR161, ENA2 (YDR039 is ENA2), and YCR105. These results imply that the regions containing these genes are extremely divergent, or all together deleted from the strain. Subsequent attempts to generate PCR products from SGE1, ENA2, and ASP3A using Y55 DNA failed. This result supports the conclusion that these genes are likely to be missing from the Y55 genome. It is interesting to note that at least two of the regions absent in the Y55 genome have been previously shown or suggested to be deleted in mutant laboratory strains (14–16). In particular, the Asp-3 region appears to be highly prone to being deleted (15, 16).

These results indicate that gene arrays can be used to efficiently screen different strains of an organism for large deletion polymorphisms. A single hybridization and scan will reveal differences based on differential hybridization to particular elements. It is reasonable to suppose that an equivalent number of genes are present in the Y55 genome and absent in the S288c genome. This result should be viewed as a minimum estimate of the deletion polymorphisms that exist between these two unrelated strains as intergenic deletions or small intragenic deletions would not be detected because considerable hybridizing material would be remain. Sequence polymorphisms, such as deletions, are present in populations of every species and must at some level affect phenotype. One of the challenges of the genome era will be to critically examine sequence polymorphisms that exist in the natural gene pool relative to the reference genome sequence.

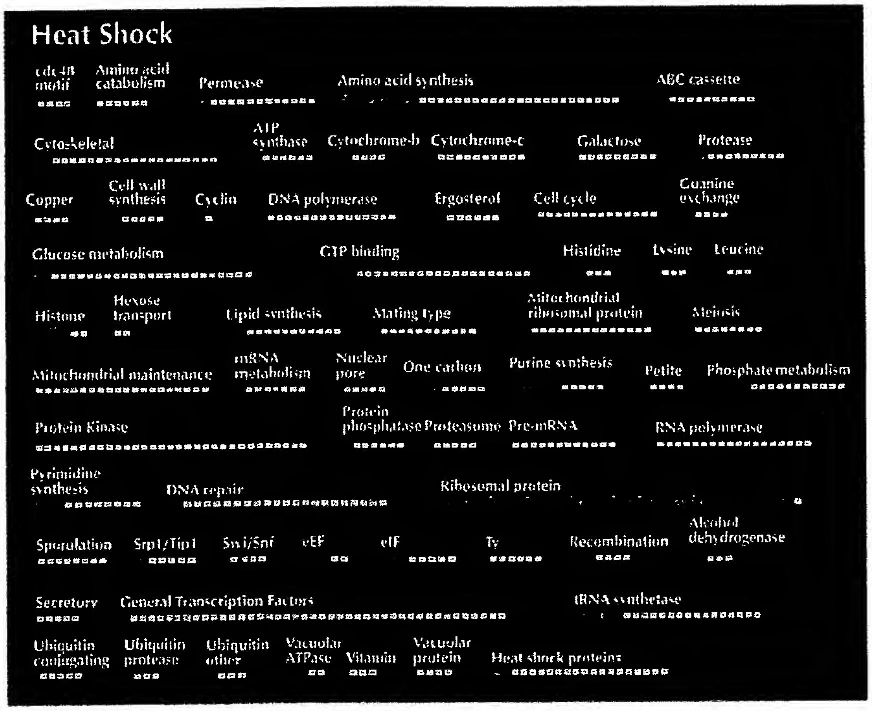


FIG. 2. ORF categories displaying differential expression between heat shocked and untreated cultures. Bars within categories correspond to individual ORFs. Green shaded bars correspond to relative increases in ORF expression under 25°C growth conditions. Red shaded bars correspond to relative increases in ORF expression under 39°C growth conditions.

Gene Expression Analysis. The arrays were used to examine gene expression in yeast grown under a variety of different conditions. Expression analysis is an ideal application of these arrays because a single hybridization provides quantitative expression data for thousands of genes. To better understand results for genes of known function, ORFs were placed in biologically relevant categories on the basis of function (e.g., amino acid catabolic genes) and/or pathways (e.g., the histidine biosynthesis pathway).

Table 2. Cold shock vs. control expression data

Ratio of gene expression				
Control	Cold	ORF	Gene	Description
	3.3	YOR153	PDR5	Pleiotropic drug resistance protein
2.4		YCR012	PGK1	Phosphoglycerate kinase
2.9		YCL040	GLK1	Aldohexose specific glucokinase
	1.4	YHR064		Heat shock protein
2.0		YJL034	KAR2	Nuclear fusion protein
2.1		YDR258	HSP78	Mitochondrial heat shock protein of clpb family of ATP-dependent proteases
2.2		YLL039	UBI4	Ubiquitin precursor
2.7		YLL026	HSP104	Heat shock protein
3.1		YER103	SSA4	Heat shock protein
3.3		YBR126	TPS1	α, α-Trehalose-phosphate synthase (UDP-forming)
3.8		YPL240	HSP82	Heat shock protein
7.9		YBR054	YRO2	Similarity to HSP30 heat shock protein Yrolp
7.9		YBR072	HSP26	Heat shock protein
16.5		YCR021	HSP30	Heat shock protein
1.8		YDR343	HXT6	High-affinity hexose transporter
2.1		YHR096	HXT5	Putative hexose transporter
2.4		YFR053	HXK1	Hexokinase I
2.8		YHR092	HXT4	Moderate- to low-affinity glucose transporter
3.4		YHR094	HXT1	Low-affinity hexose (glucose) transporter
	2.3	YHR089	GAR1	Nucleolar rRNA processing protein
	1.7	YLR048	NAB1B	40S ribosomal protein p40 homolog b
	1.7	YLR441	RP10A	Ribosomal protein S3a.e
	1.7	YLL045	RPL4B	Ribosomal protein L7a.e.B
	1.6	YLR029	RPL13A	Ribosomal protein L15.e
	1.6	YGL123	SUP44	Ribosomal protein
	3.1	YBR067	TIP1	Cold- and heat-shock-induced protein of the Srp1/Tip1p family
	2.2	YER011	TIR1	Cold-shock-induced protein of the Tir1p, Tip1p family
	2.0	YCR058		Hypothetical protein
	4.2	YKL102		Hypothetical protein

Table 3. Glucose vs. galactose expression data

Ratio of gene expression				
Glucose	Galactose	ORF	Gene	Description
2.1		YHR018	ARG4	Arginosuccinate lyase
3.5		YPR035	GLN1	Glutamate–ammonia ligase
2.8		YML116	ATR1	Aminotriazole and 4-nitroquinoline resistance protein
2.0		YMR303	ADH2	Alcohol dehydrogenase II
3.7		YBR145	ADH5	Alcohol dehydrogenase V
	3.2	YBL030	AAC2	ADP, ATP carrier protein 2
	2.9	YBR085	AAC3	ADP, ATP carrier protein
	2.7	YDR298	ATP5	H^+-transporting ATP synthase δ chain precursor
	2.5	YBR039	ATP3	H^+-transporting ATP synthase γ chain precursor
	5.5	YML054	CYB2	Lactate dehydrogenase cytochrome *b2*
	3.4	YML054	CYB2	Lactate dehydrogenase cytochrome *b2*
	2.3	YKL150	MCR1	Cytochrome-*b5* reductase
	4.2	YBL045	COR1	Ubiquinol–cytochrome *c* reductase 44K core protein
	3.5	YDL067	COX9	Cytochrome *c* oxidase chain VIIA
	2.7	YLR038	COX12	Cytochrome *c* oxidase, subunit VIB
	2.6	YHR051	COX6	Cytochrome *c* oxidase subunit VI
	2.4	YLR395	COX8	Cytochrome *c* oxidase chain VIII
	2.3	YFR033	QCR6	Ubiquinol–cytochrome *c* reductase 17K protein
	23.7	YLR081	GAL2	Galactose (and glucose) permease
	21.9	YBR018	GAL7	UDP-glucose-hexose-1-phosphate uridylyltransferase
	21.8	YBR020	GAL1	Galactokinase
	19.5	YBR019	GAL10	UDP-glucose 4-epimerase
	14.7	YLR081	GAL2	Galactose (and glucose) permease
	8.6	YDR009	GAL3	Galactokinase
	3.0	YML051	GAL80(1)	Negative regulator for expression of galactose-induced genes
	2.8	YML051	GAL80(2)	Negative regulator for expression of galactose-induced genes
2.7		YER055	HIS1	ATP phosphoribosyltransferase
3.4		YBR248	HIS7	Glutamine amidotransferase/cyclase
7.4		YCL030	HIS4	Phosphoribosyl-AMP cyclohydrolase/phosphoribosyl-ATP pyrophosphatase/histidinol dehydrogenase
5.8		YKR080	MTD1	Methylenetetrahydrofolate dehydrogenase (NAD+)
6.0		YDR019	GCV1	Glycine decarboxylase T subunit
6.1		YLR058	SHM2	Serine hydroxymethyltransferase
	8.1	YML123	PHO84	High-affinity inorganic phosphate/H^+ symporter
3.5		YDR408	ADE8	Phosphoribosylglycinamide formyltransferase (GART)
3.6		YDR408	ADE8	Phosphoribosylglycinamide formyltransferase (GART)
4.4		YAR015	ADE1	Phosphoribosylamidoimidazole-succinocarboxamide synthase
5.6		YMR300	ADE4	Amidophosphoribosyltransferase
5.6		YOR128	ADE2	Phosphoribosylaminoimidazole carboxylase
6.0		YGL234	ADE5,7	Phosphoribosylamine-glycine ligase and phosphoribosylformylglycinamidine cyclo-ligase
	6.3	YBL015	ACH1	Acetyl-CoA hydrolase

Heat Shock Results. A log phase culture growing in YEP/dextrose medium at 25°C was split in half. One half of the culture remained at 25°C whereas the other half of the culture was shifted to 39°C. mRNA was isolated from both cultures 1 h after heat shock for comparison on microarrays and, although this time point is not optimal for measuring induction of heat shock mRNAs (17), many known heat shock genes exhibited considerable induction at this time point (Table 1; Fig. 2). Down-regulation of genes in the ribosomal protein and histone gene categories was also observed. Differential expression between the heat-shocked culture and the control was also observed for many other genes. Genes in many categories, such as amino acid catabolism and amino acid synthesis, exhibited a mixed response with some genes showing little or no differential expression and other genes showing a significant increase or decrease in gene expression in response to heat shock (Table 1; Fig. 2).

Cold Shock Results. A log phase culture growing in YEP/dextrose medium at 37°C was split in half. One half of the culture remained at 37°C while the other half of the culture was shifted to 18°C. mRNA was isolated from both cultures 1 h after cold shock for comparison on microarrays. As expected, two known cold shock genes (TIP1, TIR1) were expressed at a significantly higher level in the cold-shocked culture. Genes in other functional categories, such as glucose metabolism and heat shock displayed a mixed response with expression of some genes being unaffected and other genes exhibiting significant up- or down-regulation in response to cold shock (Table 2).

Steady-State Galactose vs. Glucose Results. mRNA was isolated from steady-state log phase YEP galactose and YEP glucose grown cultures for comparison on the microarrays. As expected, the GAL genes were expressed at a much higher level in the galactose culture. Many genes were differentially expressed in these cultures that were not *a priori* expected to exhibit differential expression. For example, some genes in the amino acid catabolic category were up-regulated in the galactose culture whereas genes in the one-carbon metabolism and purine categories were largely or entirely down-regulated in the galactose culture (Table 3). Genes in other categories, such as amino acid synthesis, abc transporter, cytochrome *c*, and cytochrome *b*, exhibited mixed responses; some genes in a category showed little or no obvious differential expression whereas other genes in the same category showed significant differential expression in the galactose and glucose cultures.

DISCUSSION

The results of these experiments show that many genes are differentially expressed under the three environmental conditions described here. The expected and predicted changes in gene expression, such as HSP12 in the heat-shocked culture, TIP1 in the cold-shocked culture, and GAL2 in the steady-state galactose culture, were observed in every case. However, in addition to the expected changes in gene expression, significant differential expression was also observed for many other genes that would not, *a priori*, be expected to be differentially expressed. For example, expression of PHO11 decreased and expression of YLR194, KIN2, and HXT6 increased in the heat shocked culture. Expression of MST1 and APE3 decreased and expression of PDR5 and GAR1 increased in the cold-shocked culture. In addition, ADE4 and SER2 were expressed at reduced levels whereas PHO84 and ACH1 were expressed at higher levels in cells grown in galactose compared with cells grown in glucose. Differential expression of these and many other genes was specific to one of these three environmental conditions.

Many other genes were found to be differentially expressed under more than one condition. When differentially expressed genes in cold- and heat-shocked cultures were compared, 30 genes were found in common. Of these 30 genes, 28 showed inverse expression (i.e., increased expression under one condition and decreased expression under the other condition). Two genes, YCR058 and YKL102, showed elevated expression in response to both cold and heat shock. Fifteen genes were found to be differentially expressed in both the heat-shocked and steady-state galactose cultures: 9 genes showed increased expression and 5 showed decreased expression under both conditions. Twenty genes were differentially expressed in both the cold-shocked and steady-state galactose cultures: 8 genes showed decreased expression and 5 genes showed increased expression under both conditions. Six genes showed increased expression in the galactose culture and decreased expression in the cold shocked culture. One gene (ODP1) showed increased expression in both the cold-shocked and steady-state galactose cultures.

Gene expression is affected in a global fashion when environmental conditions are changed and both expected and unexpected genes are affected. There is also overlap in the genes that are differentially expressed under quite different environmental conditions. These results can be rationalized by considering the high degree of cross-pathway regulation in yeast. For example, there is evidence for cross-pathway regulation between (*i*) carbon and nitrogen metabolism (18), (*ii*) phosphate and sulfate metabolism (19), and (*iii*) purine, phosphate, and amino acid metabolism (20–24). There are also examples of the interaction of general and specific transcription factors (25, 26). Finally, within the broad class of amino acid biosynthetic genes, there is evidence for amino acid specific regulation of some genes, regulation via general control for other genes, and regulation via both specific and general control for other genes (22, 27–30).

Cross-pathway regulation arises from the complex structure of promoters. Virtually all promoters contain sites for multiple transcription factors and, therefore, virtually all genes are subject to combinatorial regulation. For example, the HIS4 promoter contains binding sites for GCN4 (the general amino acid control transcription factor), PHO2/BAS2 (a transcriptional regulator of phosphatase and purine biosynthetic genes), and BAS1 (a transcriptional regulator of purine biosynthetic genes) (31). It is likely that the complex effects on gene expression described in this work are a direct consequence of the combinatorial regulation of gene expression.

These findings illustrate the power of the highly parallel whole genome approach when examining gene expression. The global effects of environmental change on gene expression can now be directly visualized. It is clear that determining the mechanism(s) and the functional role of the dramatic global effects on gene expression in different environments will be a significant challenge. The era of whole genome analysis will, ultimately, allow researchers to switch from the very focused single gene/promoter view of gene expression and instead view the cell more as a large complex network of gene regulatory pathways.

With the entire sequence of this model organism known, new approaches have been developed that allow for genome wide analyses (32, 33) of gene function. The genome microarrays represent a novel tool for genetic and expression analysis of the yeast genome. This pilot study uses arrays containing >35% of the yeast ORFs and it is clear that the entire set of ORFs from the yeast genome can be arrayed using the directed primer based strategy detailed here. Recent advances in arraying technology will allow all 6,100 ORFs to be arrayed in an area of less than 1.8 cm^2. Furthermore, as the technology improves, detection limits will allow less than 500 ng of starting mRNA material to be used for making probe.

The genome arrays provide for a robust, fully automated approach toward examining genome structure and gene function. They allow for comparisons between different genomes as well as a detailed study of gene expression at the global level. This research will help to elucidate relationships between genes and allow the researcher to understand gene function by understanding expression patterns across the yeast genome.

Support was provided by National Institutes of Health Grant P0/HG00205.

1. Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., *et al.* (1995) *Science* 269, 496–512.
2. Fraser, C. M., Gocayne, J. D., White, O., Adams, M. D., Clayton, R. A., *et al.* (1995) *Science* 270, 397–403.
3. Bult, C. J., White, O., Olsen, G. J., Zhou, L., Fleischmann, R. D., *et al.* (1996) *Science* 273, 1058–1073.
4. Sulston, J., Du, Z., Thomas, K., Wilson, R., Hillier, L., *et al.* (1992) *Nature (London)* 356, 37.
5. Newman, T., de Bruijn, F. J., Green, P., Keegstra, K., Kende, H., *et al.* (1994) *Plant Physiol.* 106, 1241–1255.
6. Lashkari, D. A., Hunicke-Smith, S. P., Norgren, R. M., Davis, R. W. & Brennan, T. (1995) *Proc. Natl. Acad. Sci. USA* 92, 7912–7915.
7. Schena, M., Shalon, D., Davis, R. W. & Brown, P. O. (1995) *Science* 270, 467–470.
8. Shalon, D., Smith, S. & Brown, P. O. (1996) *Genome Res.* 6, 639–645.
9. Heller, R. A., Schena, M., Chai, A., Shalon, D., Bedilion, T., Gilmore, J., Woolley, D. E. & Davis, R. W. (1997) *Proc. Natl. Acad. Sci. USA* 94, 2150–2155.
10. DeRisi, J., Penland, L., brown, P. O., Bittner, M. L., Meltzer, P. S., Ray, M., Chen, Y., Su Ya & Trent, J. M. (1996) *Nat. Genet.* 14, 457–460.
11. Nelson, S. F., McCusker, J. H., Sander, M. A., Kee, Y., Modrich, P. & Brown P. O. (1993) *Nat. Genet.* 4, 11–18.
12. Hoffman, C. S. & Winston, F. (1989) *Gene* 84, 473–479.
13. Schmitt, M., Brown, T. & Trumpower, B. (1990) *Nucleic Acids Res.* 18, 3091.
14. Ehrenhofer-Murray, A. E., Wurgler, F. E. & Sengstag, C. (1994) *Mol. Gen. Genet.* 244, 287–294.
15. Kim, K-W., Kamerud, J. Q., Livingston, D. M. & Roon, R. J. (1988) *J. Biol. Chem.* 263, 11948–11953.
16. Kim, K.-W. & Roon, R. J. (1984) *J. Bacteriol.* 157, 958–961.
17. Craig, E. A. (1992) in *The Molecular Biology of the Yeast Saccharomyces: Gene Expression*, eds. Jones, E. W., Pringle, J. R. & Broach, J. R. (Cold Spring Harbor Lab. Press, Plainview, NY), Vol. 2, pp. 501–537.
18. Dang, V. D., Bohn, C., Bolotin-Fukuhara, M. & Daignan-Fornier, B. (1996) *J. Bacteriol.* 178, 1842–1849.
19. O'Connell, K. F. & Baker, R. E. (1992) *Genetics* 132, 63–73.
20. Braus, G., Mosch, H. U., Vogel, K., Hinnen, A. & Hutter, R. (1989) *EMBO J.* 8, 939–945.
21. Mosch, H. U., Scheier, B., Lahti, R., Mantsala, P. & Braus, G. H. (1991) *J. Biol. Chem.* 266, 20453–20456.
22. Mitchell, A. P. & Magasanik, B. (1984) *Mol. Cell. Biol.* 4, 2767–2773.
23. Daignan-Fornier, B. & Fink, G. R. (1992) *Proc. Natl. Acad. Sci USA* 89, 6746–6750.
24. Tice-Baldwin, K., Fink, G. R. & Arndt, K. T. (1989) *Science* 246, 931–935.
25. Messenguy, F. & Dubois, E. (1993) *Mol. Cell. Biol.* 13, 2586–2592.
26. Devlin, C., Tice-Baldwin, K., Shore, D. & Arndt, K. T. (1991) *Mol. Cell. Biol.* 11, 3642–3651.
27. Magasanik, B. (1992) in *The Molecular and Cellular Biology of the Yeast Saccharomyces: Gene Expression*, eds. Jones, E. W., Pringle, J. R. & Broach, J. R. (Cold Spring Harbor Lab. Press, Plainview, NY), Vol. 2, pp. 283–317.
28. Hinnebusch, A. G. (1992) in *The Molecular and Cellular Biology of the Yeast Saccharomyces: Gene Expression*, eds. Jones, E. W., Pringle, J. R. & Broach, J. R. (Cold Spring Harbor Lab. Press, Plainview, NY), Vol. 2, pp. 319–414.
29. Brisco, P. R. & Kohlhaw, G. B. (1990) *J. Biol. Chem.* 265, 11667–11675.
30. O'Connell, K. F., Surdin-Kerjan, Y. & Baker R. E. (1995) *Mol. Cell. Biol.* 15, 1879–1888.
31. Arndt K. T., Styles, C. & Fink, G. R. (1987) *Science* 237, 874–880.
32. Smith, V., Chou, K. N., Lashkari, D., Botstein, D. & Brown, P. O. (1996) *Science* 274, 2069–2074.
33. Shoemaker, D. D., Lashkari, D. A., Morris, D., Mittman, M. & Davis, R. W. (1996) *Nat. Genet.* 14, 450–456.

Fischer-Vize, *Science* 270, 1828 (1995).
35. T. C. James and S. C. Elgin, *Mol. Cell Biol.* 6, 3862 (1986); R. Paro and D. S. Hogness, *Proc. Natl. Acad. Sci. U.S.A.* 88, 263 (1991); B. Tschiersch *et al.*, *EMBO J.* 13, 3822 (1994); M. T. Madireddi *et al.*, *Cell* 87, 75 (1996); D. G. Stokes, K. D. Tartof, R. P. Perry, *Proc. Natl. Acad. Sci. U.S.A.* 93, 7137 (1996).
36. P. M. Palosaari *et al.*, *J. Biol. Chem.* 266, 10750 (1991); A. Schmitz, K. H. Gartemann, J. Fiedler, E. Grund, R. Eichenlaub, *Appl. Environ. Microbiol.* 58, 4068 (1992); V. Sharma, K. Suvarna, R. Meganathan, M. E. Hudspeth, *J. Bacteriol.* 174, 5057 (1992); M. Kanazawa *et al.*, *Enzyme Protein* 47, 9 (1993); Z. L. Boynton, G. N. Bennet, F. B. Rudolph, *J. Bacteriol.* 178, 3015 (1996).
37. M. Ho *et al.*, *Cell* 77, 869 (1994).
38. W. Hendriks *et al.*, *J. Cell Biochem.* 59, 418 (1995).
39. We thank H. Skaletsky and F. Lewitter for help with sequence analysis; Lawrence Livermore National Laboratory for the flow-sorted Y cosmid library; and P. Bain, A. Bortvin, A. de la Chapelle, G. Fink, K. Jegalian, T. Kawaguchi, E. Lander, H. Lodish, P. Matsudaira, D. Menke, U. RajBhandary, R. Reijo, S. Rozen, A. Schwartz, C. Sun, and C. Tilford for comments on the manuscript. Supported by NIH.

28 April 1997; accepted 9 September 1997

Exploring the Metabolic and Genetic Control of Gene Expression on a Genomic Scale

Joseph L. DeRisi, Vishwanath R. Iyer, Patrick O. Brown*

DNA microarrays containing virtually every gene of *Saccharomyces cerevisiae* were used to carry out a comprehensive investigation of the temporal program of gene expression accompanying the metabolic shift from fermentation to respiration. The expression profiles observed for genes with known metabolic functions pointed to features of the metabolic reprogramming that occur during the diauxic shift, and the expression patterns of many previously uncharacterized genes provided clues to their possible functions. The same DNA microarrays were also used to identify genes whose expression was affected by deletion of the transcriptional co-repressor *TUP1* or overexpression of the transcriptional activator *YAP1*. These results demonstrate the feasibility and utility of this approach to genomewide exploration of gene expression patterns.

The complete sequences of nearly a dozen microbial genomes are known, and in the next several years we expect to know the complete genome sequences of several metazoans, including the human genome. Defining the role of each gene in these genomes will be a formidable task, and understanding how the genome functions as a whole in the complex natural history of a living organism presents an even greater challenge.

Knowing when and where a gene is expressed often provides a strong clue as to its biological role. Conversely, the pattern of genes expressed in a cell can provide detailed information about its state. Although regulation of protein abundance in a cell is by no means accomplished solely by regulation of mRNA, virtually all differences in cell type or state are correlated with changes in the mRNA levels of many genes. This is fortuitous because the only specific reagent required to measure the abundance of the mRNA for a specific gene is a cDNA sequence. DNA microarrays, consisting of thousands of individual gene sequences printed in a high-density array on a glass microscope slide (*1*, *2*), provide a practical and economical tool for studying gene expression on a very large scale (3–6).

Saccharomyces cerevisiae is an especially favorable organism in which to conduct a systematic investigation of gene expression. The genes are easy to recognize in the genome sequence, *cis* regulatory elements are generally compact and close to the transcription units, much is already known about its genetic regulatory mechanisms, and a powerful set of tools is available for its analysis.

A recurring cycle in the natural history of yeast involves a shift from anaerobic (fermentation) to aerobic (respiration) metabolism. Inoculation of yeast into a medium rich in sugar is followed by rapid growth fueled by fermentation, with the production of ethanol. When the fermentable sugar is exhausted, the yeast cells turn to ethanol as a carbon source for aerobic growth. This switch from anaerobic growth to aerobic respiration upon depletion of glucose, referred to as the diauxic shift, is correlated with widespread changes in the expression of genes involved in fundamental cellular processes such as carbon metabolism, protein synthesis, and carbohydrate storage (*7*). We used DNA microarrays to characterize the changes in gene expression that take place during this process for nearly the entire genome, and to investigate the genetic circuitry that regulates and executes this program.

Yeast open reading frames (ORFs) were amplified by the polymerase chain reaction (PCR), with a commercially available set of primer pairs (*8*). DNA microarrays, containing approximately 6400 distinct DNA sequences, were printed onto glass slides by using a simple robotic printing device (*9*). Cells from an exponentially growing culture of yeast were inoculated into fresh medium and grown at 30°C for 21 hours. After an initial 9 hours of growth, samples were harvested at seven successive 2-hour intervals, and mRNA was isolated (*10*). Fluorescently labeled cDNA was prepared by reverse transcription in the presence of Cy3(green)- or Cy5(red)-labeled deoxyuridine triphosphate (dUTP) (*11*) and then hybridized to the microarrays (*12*). To maximize the reliability with which changes in expression levels could be discerned, we labeled cDNA prepared from cells at each successive time point with Cy5, then mixed it with a Cy3-labeled "reference" cDNA sample prepared from cells harvested at the first interval after inoculation. In this experimental design, the relative fluorescence intensity measured for the Cy3 and Cy5 fluors at each array element provides a reliable measure of the relative abundance of the corresponding mRNA in the two cell populations (Fig. 1). Data from the series of seven samples (Fig. 2), consisting of more than 43,000 expression-ratio measurements, were organized into a database to facilitate efficient exploration and analysis of the results. This database is publicly available on the Internet (*13*).

During exponential growth in glucose-rich medium, the global pattern of gene expression was remarkably stable. Indeed, when gene expression patterns between the first two cell samples (harvested at a 2-hour interval) were compared, mRNA levels differed by a factor of 2 or more for only 19 genes (0.3%), and the largest of these differences was only 2.7-fold (*14*). However, as glucose was progressively depleted from the growth media during the course of the experiment, a marked change was seen in the global pattern of gene expression. mRNA levels for approximately 710 genes were induced by a factor of at least 2, and the mRNA levels for approximately 1030 genes declined by a factor of at least 2. Messenger RNA levels for 183 genes increased by a factor of at least 4, and mRNA levels for 203 genes diminished by a factor of at least 4. About half of these differentially expressed genes have no currently recognized function and are not yet named. Indeed, more than 400 of the differentially expressed genes have no apparent homology

Department of Biochemistry, Stanford University School of Medicine, Howard Hughes Medical Institute, Stanford, CA 94305–5428, USA.

*To whom correspondence should be addressed. E-mail: pbrown@cmgm.stanford.edu

to any gene whose function is known (*15*). The responses of these previously uncharacterized genes to the diauxic shift therefore provides the first small clue to their possible roles.

The global view of changes in expression of genes with known functions provides a vivid picture of the way in which the cell adapts to a changing environment. Figure 3 shows a portion of the yeast metabolic pathways involved in carbon and energy metabolism. Mapping the changes we observed in the mRNAs encoding each enzyme onto this framework allowed us to infer the redirection in the flow of metabolites through this system. We observed large inductions of the genes coding for the enzymes aldehyde dehydrogenase (*ALD2*) and acetyl-coenzyme A(CoA) synthase (*ACS1*), which function together to convert the products of alcohol dehydrogenase into acetyl-CoA, which in turn is used to fuel the tricarboxylic acid (TCA) cycle and the glyoxylate cycle. The concomitant shutdown of transcription of the genes encoding pyruvate decarboxylase and induction of pyruvate carboxylase rechannels pyruvate away from acetaldehyde, and instead to oxalacetate, where it can serve to supply the TCA cycle and gluconeogenesis. Induction of the pivotal genes *PCK1*, encoding phosphoenolpyruvate carboxykinase, and *FBP1*, encoding fructose 1,6-biphosphatase, switches the directions of two key irreversible steps in glycolysis, reversing the flow of metabolites along the reversible steps of the glycolytic pathway toward the essential biosynthetic precursor, glucose-6-phosphate. Induction of the genes coding for the trehalose synthase and glycogen synthase complexes promotes channeling of glucose-6-phosphate into these carbohydrate storage pathways.

Just as the changes in expression of genes encoding pivotal enzymes can provide insight into metabolic reprogramming, the behavior of large groups of functionally related genes can provide a broad view of the systematic way in which the yeast cell adapts to a changing environment (Fig. 4). Several classes of genes, such as cytochrome c–related genes and those involved in the TCA/glyoxylate cycle and carbohydrate storage, were coordinately induced by glucose exhaustion. In contrast, genes devoted to protein synthesis, including ribosomal proteins, tRNA synthetases, and translation, elongation, and initiation factors, exhibited a coordinated decrease in expression. More than 95% of ribosomal genes showed at least twofold decreases in expression during the diauxic shift (Fig. 4) (*13*). A noteworthy and illuminating exception was that the genes encoding mitochondrial ribosomal genes were generally induced rather than repressed after glucose limitation, highlighting the requirement for mitchondrial biogenesis (*13*). As more is learned about the functions of every gene in the yeast genome, the ability to gain insight into a cell's response to a changing environment through its global gene expression patterns will become increasingly powerful.

Several distinct temporal patterns of expression could be recognized, and sets of genes could be grouped on the basis of the similarities in their expression patterns. The characterized members of each of these groups also shared important similarities in their functions. Moreover, in most cases, common regulatory mechanisms could be inferred for sets of genes with similar expression profiles. For example, seven genes showed a late induction profile, with mRNA levels increasing by more than ninefold at the last timepoint but less than threefold at the preceding timepoint (Fig. 5B). All of these genes were known to be glucose-repressed, and five of the seven were previously noted to share a common upstream activating sequence (UAS), the carbon source response element (CSRE) (*16*–*20*). A search in the promoter regions of the remaining two genes, *ACR1* and *IDP2*, revealed that *ACR1*, a gene essential for *ACS1* activity, also possessed a consensus CSRE motif, but interestingly, *IDP2* did not. A search of the entire yeast genome sequence for the consensus CSRE motif revealed only four additional candidate genes, none of which showed a similar induction.

Examples from additional groups of genes that shared expression profiles are illustrated in Fig. 5, C through F. The sequences upstream of the named genes in Fig. 5C all contain stress response elements (STRE), and with the exception

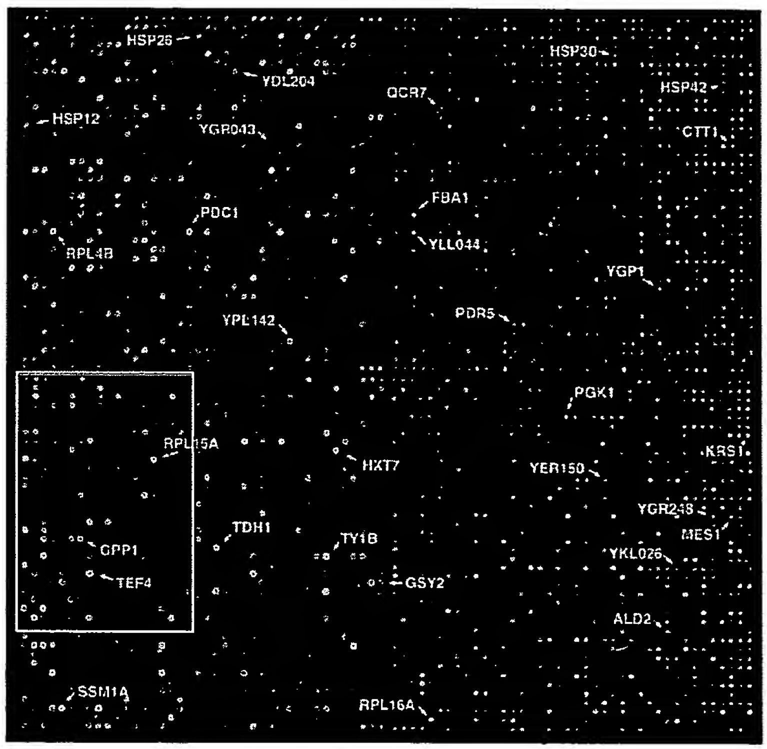


Fig. 1. Yeast genome microarray. The actual size of the microarray is 18 mm by 18 mm. The microarray was printed as described (*9*). This image was obtained with the same fluorescent scanning confocal microscope used to collect all the data we report (*49*). A fluorescently labeled cDNA probe was prepared from mRNA isolated from cells harvested shortly after inoculation (culture density of $<5 \times 10^6$ cells/ml and media glucose level of 19 g/liter) by reverse transcription in the presence of Cy3-dUTP. Similarly, a second probe was prepared from mRNA isolated from cells taken from the same culture 9.5 hours later (culture density of $\sim 2 \times 10^8$ cells/ml, with a glucose level of <0.2 g/liter) by reverse transcription in the presence of Cy5-dUTP. In this image, hybridization of the Cy3-dUTP–labeled cDNA (that is, mRNA expression at the initial timepoint) is represented as a green signal, and hybridization of Cy5-dUTP–labeled cDNA (that is, mRNA expression at 9.5 hours) is represented as a red signal. Thus, genes induced or repressed after the diauxic shift appear in this image as red and green spots, respectively. Genes expressed at roughly equal levels before and after the diauxic shift appear in this image as yellow spots.

of *HSP42*, have previously been shown to be controlled at least in part by these elements (*21–24*). Inspection of the sequences upstream of *HSP42* and the two uncharacterized genes shown in Fig. 5C, YKL026c, a hypothetical protein with similarity to glutathione peroxidase, and YGR043c, a putative transaldolase, revealed that each of these genes also possess repeated upstream copies of the stress-responsive CCCCT motif. Of the 13 additional genes in the yeast genome that shared this expression profile [including *HSP30*, *ALD2*, *OM45*, and 10 uncharacterized ORFs (*25*)], nine contained one or more recognizable STRE sites in their upstream regions.

The heterotrimeric transcriptional activator complex *HAP2,3,4* has been shown to be responsible for induction of several genes important for respiration (*26–28*). This complex binds a degenerate consensus sequence known as the CCAAT box (*26*). Computer analysis, using the consensus sequence TNRYTGGB (*29*), has suggested that a large number of genes involved in respiration may be specific targets of *HAP2,3,4* (*30*). Indeed, a putative *HAP2,3,4* binding site could be found in the sequences upstream of each of the seven cytochrome c–related genes that showed the greatest magnitude of induction (Fig. 5D). Of 12 additional cytochrome c–related genes that were induced, *HAP2,3,4* binding sites were present in all but one. Significantly, we found that transcription of *HAP4* itself was induced nearly ninefold concomitant with the diauxic shift.

Control of ribosomal protein biogenesis is mainly exerted at the transcriptional level, through the presence of a common upstream-activating element (UAS_{rpg}) that is recognized by the Rap1 DNA-binding protein (*31*, *32*). The expression profiles of seven ribosomal proteins are shown in Fig. 5F. A search of the sequences upstream of all seven genes revealed consensus Rap1-binding motifs (*33*). It has been suggested that declining Rap1 levels in the cell during starvation may be responsible for the decline in ribosomal protein gene expression (*34*). Indeed, we observed that the abundance of *RAP1* mRNA diminished by 4.4-fold, at about the time of glucose exhaustion.

Of the 149 genes that encode known or putative transcription factors, only two, *HAP4* and *SIP4*, were induced by a factor of more than threefold at the diauxic shift. *SIP4* encodes a DNA-binding transcriptional activator that has been shown to interact with Snf1, the "master regulator" of glucose repression (*35*). The eightfold induction of *SIP4* upon depletion of glucose strongly suggests a role in the induction of downstream genes at the diauxic shift.

Although most of the transcriptional responses that we observed were not previously known, the responses of many genes during the diauxic shift have been described. Comparison of the results we obtained by DNA microarray hybridization with previously reported results therefore provided a strong test of the sensitivity and accuracy of this approach. The expression patterns we observed for previously characterized genes showed almost perfect concordance with previously published results (*36*). Moreover, the differential expression measurements obtained by DNA microarray hybridization were reproducible in duplicate experiments. For example, the remarkable changes in gene expression between cells harvested immediately after inoculation and immediately after the diauxic shift (the first and sixth intervals in this time series) were measured in duplicate, independent DNA microarray hybridizations. The correlation coefficient for two complete sets of expression ratio measurements was 0.87, and for more than 95% of the genes, the expression ratios measured in these duplicate experiments differed by less than a factor of 2. However, in a few cases, there were discrepancies between our results and previous results, pointing to technical limitations that will need to be addressed as DNA microarray technology advances (*37*, *38*). Despite the noted exceptions, the high concordance between the results we obtained in these experiments and those of previous studies provides confidence in the reliability and thoroughness of the survey.

The changes in gene expression during this diauxic shift are complex and involve integration of many kinds of information about the nutritional and metabolic state of the cell. The large number of genes whose expression is altered and the diversity of temporal expression profiles observed in this experiment highlight the challenge of understanding the underlying regulatory mechanisms. One approach to defining the contributions of individual regulatory genes to a complex program of this kind is to use DNA microarrays to identify genes whose expression is affected

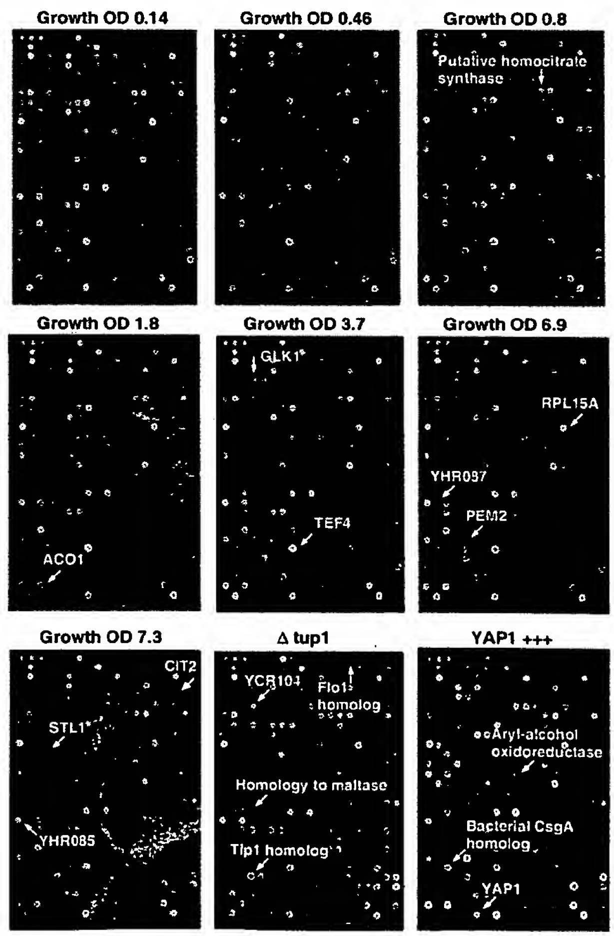


Fig. 2. The section of the array indicated by the gray box in Fig. 1 is shown for each of the experiments described here. Representative genes are labeled. In each of the arrays used to analyze gene expression during the diauxic shift, red spots represent genes that were induced relative to the initial timepoint, and green spots represent genes that were repressed relative to the initial timepoint. In the arrays used to analyze the effects of the *tup1*Δ mutation and *YAP1* overexpression, red spots represent genes whose expression was increased, and green spots represent genes whose expression was decreased by the genetic modification. Note that distinct sets of genes are induced and repressed in the different experiments. The complete images of each of these arrays can be viewed on the Internet (*13*). Cell density as measured by optical density (OD) at 600 nm was used to measure the growth of the culture.

by mutations in each putative regulatory gene. As a test of this strategy, we analyzed the genomewide changes in gene expression that result from deletion of the *TUP1* gene. Transcriptional repression of many genes by glucose requires the DNA-binding repressor Mig1 and is mediated by recruiting the transcriptional co-repressors Tup1 and Cyc8/Ssn6 (39). Tup1 has also been implicated in repression of oxygen-regulated, mating-type–specific, and DNA-damage–inducible genes (40).

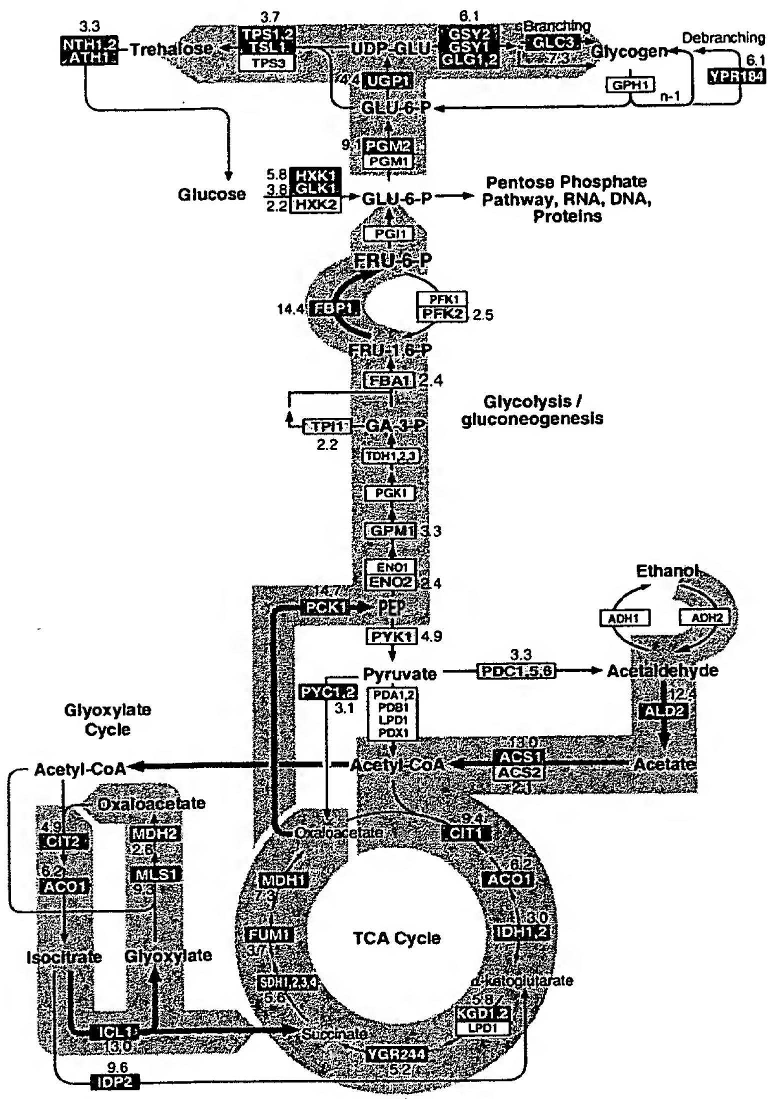


Fig. 3. Metabolic reprogramming inferred from global analysis of changes in gene expression. Only key metabolic intermediates are identified. The yeast genes encoding the enzymes that catalyze each step in this metabolic circuit are identified by name in the boxes. The genes encoding succinyl-CoA synthase and glycogen-debranching enzyme have not been explicitly identified, but the ORFs YGR244 and YPR184 show significant homology to known succinyl-CoA synthase and glycogen-debranching enzymes, respectively, and are therefore included in the corresponding steps in this figure. Red boxes with white lettering identify genes whose expression increases in the diauxic shift. Green boxes with dark green lettering identify genes whose expression diminishes in the diauxic shift. The magnitude of induction or repression is indicated for these genes. For multimeric enzyme complexes, such as succinate dehydrogenase, the indicated fold-induction represents an unweighted average of all the genes listed in the box. Black and white boxes indicate no significant differential expression (less than twofold). The direction of the arrows connecting reversible enzymatic steps indicate the direction of the flow of metabolic intermediates, inferred from the gene expression pattern, after the diauxic shift. Arrows representing steps catalyzed by genes whose expression was strongly induced are highlighted in red. The broad gray arrows represent major increases in the flow of metabolites after the diauxic shift, inferred from the indicated changes in gene expression.

Wild-type yeast cells and cells bearing a deletion of the *TUP1* gene (*tup1*Δ) were grown in parallel cultures in rich medium containing glucose as the carbon source. Messenger RNA was isolated from exponentially growing cells from the two populations and used to prepare cDNA labeled with Cy3 (green) and Cy5 (red), respectively (*11*). The labeled probes were mixed and simultaneously hybridized to the microarray. Red spots on the microarray therefore represented genes whose transcription was induced in the *tup1*Δ strain, and thus presumably repressed by Tup1 (*41*). A representative section of the microarray (Fig. 2, bottom middle panel) illustrates that the genes whose expression was affected by the *tup1*Δ mutation, were, in general, distinct from those induced upon glucose exhaustion [complete images of all the arrays shown in Fig. 2 are available on the Internet (*13*)]. Nevertheless, 34 (10%) of the genes that were induced by a factor of at least 2 after the diauxic shift were similarly induced by deletion of *TUP1*, suggesting that these genes may be subject to *TUP1*-mediated repression by glucose. For example, *SUC2*, the gene encoding invertase, and all five hexose transporter genes that were induced during the course of the diauxic shift were similarly induced, in duplicate experiments, by the deletion of *TUP1*.

The set of genes affected by Tup1 in this experiment also included α-glucosidases, the mating-type–specific genes *MFA1* and *MFA2*, and the DNA damage–inducible *RNR2* and *RNR4*, as well as genes involved in flocculation and many genes of unknown function. The hybridization signal corresponding to expression of *TUP1* itself was also severely reduced because of the (incomplete) deletion of the transcription unit in the *tup1*Δ strain, providing a positive control in the experiment (*42*).

Many of the transcriptional targets of Tup1 fell into sets of genes with related biochemical functions. For instance, although only about 3% of all yeast genes appeared to be *TUP1*-repressed by a factor of more than 2 in duplicate experiments under these conditions, 6 of the 13 genes that have been implicated in flocculation (*15*) showed a reproducible increase in expression of at least twofold when *TUP1* was deleted. Another group of related genes that appeared to be subject to *TUP1* repression encodes the serine-rich cell wall mannoproteins, such as Tip1 and Tir1/Srp1 which are induced by cold shock and other stresses (*43*), and similar, serine-poor proteins, the seripauperins (*44*). Messenger RNA levels for 23 of the 26 genes in this group were reproducibly elevated by at least 2.5-fold in the *tup1*Δ

strain, and 18 of these genes were induced by more than sevenfold when *TUP1* was deleted. In contrast, none of 83 genes that could be classified as putative regulators of the cell division cycle were induced more than twofold by deletion of *TUP1*. Thus, despite the diversity of the regulatory systems that employ Tup1, most of the genes that it regulates under these conditions fall into a limited number of distinct functional classes.

Because the microarray allows us to monitor expression of nearly every gene in yeast, we can, in principle, use this approach to identify all the transcriptional targets of a regulatory protein like Tup1. It is important to note, however, that in any single experiment of this kind we can only recognize those target genes that are normally repressed (or induced) under the conditions of the experiment. For instance, the experiment described here analyzed a *MAT* α strain in which *MFA1* and *MFA2*, the genes encoding the a-factor mating pheromone precursor, are normally repressed. In the isogenic *tup1Δ* strain, these genes were inappropriately expressed, reflecting the role that Tup1 plays in their repression. Had we instead carried out this experiment with a *MATA* strain (in which expression of *MFA1* and *MFA2* is not repressed), it would not have been possible to conclude anything regarding the role of Tup1 in the repression of these genes. Conversely, we cannot distinguish indirect effects of the chronic absence of Tup1 in the mutant strain from effects directly attributable to its participation in repressing the transcription of a gene.

Another simple route to modulating the activity of a regulatory factor is to overexpress the gene that encodes it. *YAP1* encodes a DNA-binding transcription factor belonging to the b-zip class of DNA-binding proteins. Overexpression of *YAP1* in yeast confers increased resistance to hydrogen peroxide, *o*-phenanthroline, heavy metals, and osmotic stress (*45*). We analyzed differential gene expression between a wild-type strain bearing a control plasmid and a strain with a plasmid expressing *YAP1* under the control of the strong *GAL1-10* promoter, both grown in galactose (that is, a condition that induces *YAP1* overexpression). Complementary DNA from the control and *YAP1* overexpressing strains, labeled with Cy3 and Cy5, respectively, was prepared from mRNA isolated from the two strains and hybridized to the microarray. Thus, red spots on the array represent genes that were induced in the strain overexpressing *YAP1*.

Of the 17 genes whose mRNA levels increased by more than threefold when *YAP1* was overexpressed in this way, five bear homology to aryl-alcohol oxidoreductases (Fig. 2 and Table 1). An additional four of the genes in this set also belong to the general class of dehydrogenases/oxidoreductases. Very little is known about the role of aryl-alcohol oxidoreductases in *S. cerevisiae*, but these enzymes have been isolated from ligninolytic fungi, in which they participate in coupled redox reactions, oxidizing aromatic, and aliphatic unsaturated alcohols to aldehydes with the production of hydrogen peroxide (*46, 47*). The fact that a remarkable fraction of the targets identified in this experiment belong to the same small, functional group of oxidoreductases suggests that these genes might play an important protective role during oxidative stress. Transcription of a small number of genes was reduced in the strain overexpressing Yap1. Interestingly, many of these genes encode sugar permeases or enzymes involved in inositol metabolism.

We searched for Yap1-binding sites (TTACTAA or TGACTAA) in the sequences upstream of the target genes we identified (*48*). About two-thirds of the genes that were induced by more than threefold upon Yap1 overexpression had one or more binding sites within 600 bases upstream of the start codon (Table 1), suggesting that they are directly regulated by Yap1. The absence of canonical Yap1-bind-

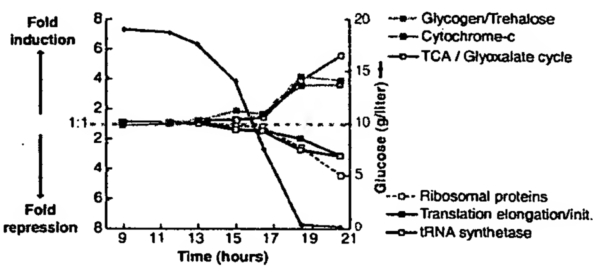


Fig. 4. Coordinated regulation of functionally related genes. The curves represent the average induction or repression ratios for all the genes in each indicated group. The total number of genes in each group was as follows: ribosomal proteins, 112; translation elongation and initiation factors, 25; tRNA synthetases (excluding mitochondial synthetases), 17; glycogen and trehalose synthesis and degradation, 15; cytochrome c oxidase and reductase proteins, 19; and TCA- and glyoxylate-cycle enzymes, 24.

Table 1. Genes induced by *YAP1* overexpression. This list includes all the genes for which mRNA levels increased by more than twofold upon *YAP1* overexpression in both of two duplicate experiments, and for which the average increase in mRNA level in the two experiments was greater than threefold (*50*). Positions of the canonical Yap1 binding sites upstream of the start codon, when present, and the average fold-increase in mRNA levels measured in the two experiments are indicated.

ORF	Distance of Yap1 site from ATG	Gene	Description	Fold-increase
YNL331C			Putative aryl-alcohol reductase	12.9
YKL071W	162–222 (5 sites)		Similarity to bacterial csgA protein	10.4
YML007W		YAP1	Transcriptional activator involved in oxidative stress response	9.8
YFL056C	223, 242		Homology to aryl-alcohol dehydrogenases	9.0
YLL060C	98		Putative glutathione transferase	7.4
YOL165C	266		Putative aryl-alcohol dehydrogenase (NADP+)	7.0
YCR107W			Putative aryl-alcohol reductase	6.5
YML116W	409	ATR1	Aminotriazole and 4-nitroquinoline resistance protein	6.5
YBR008C	142, 167, 364		Homology to benomyl/methotrexate resistance protein	6.1
YCLX08C			Hypothetical protein	6.1
YJR155W			Putative aryl-alcohol dehydrogenase	6.0
YPL171C	148, 212	OYE3	NAPDH dehydrogenase (old yellow enzyme), isoform 3	5.8
YLR460C	167, 317		Homology to hypothetical proteins YCR102c and YNL134c	4.7
YKR076W	178		Homology to hypothetical protein YMR251w	4.5
YHR179W	327	OYE2	NAD(P)H oxidoreductase (old yellow enzyme), isoform 1	4.1
YML131W	507		Similarity to *A. thaliana* zeta-crystallin homolog	3.7
YOL126C		MDH2	Malate dehydrogenase	3.3

ing sites upstream of the others may reflect an ability of Yap1 to bind sites that differ from the canonical binding sites, perhaps in cooperation with other factors, or less likely, may represent an indirect effect of Yap1 overexpression, mediated by one or more intermediary factors. Yap1 sites were found only four times in the corresponding region of an arbitrary set of 30 genes that were not differentially regulated by Yap1.

Use of a DNA microarray to characterize the transcriptional consequences of mutations affecting the activity of regulatory molecules provides a simple and powerful approach to dissection and characterization of regulatory pathways and networks. This strategy also has an important practical application in drug screening. Mutations in specific genes encoding candidate drug targets can serve as surrogates for the ideal chemical inhibitor or modulator of their activity. DNA microarrays can be used to define the resulting signature pattern of alterations in gene expression, and then subsequently used in an assay to screen for compounds that reproduce the desired signature pattern.

DNA microarrays provide a simple and economical way to explore gene expression patterns on a genomic scale. The hurdles to extending this approach to any other organism are minor. The equipment required for fabricating and using DNA microarrays (9) consists of components that were chosen for their modest cost and simplicity. It was feasible for a small group to accomplish the amplification of more than 6000 genes in about 4 months and, once the amplified gene sequences were in hand, only 2 days were required to print a set of 110 microarrays of 6400 elements each. Probe preparation, hybridization, and fluorescent imaging are also simple procedures. Even conceptually simple experiments, as we described here, can yield vast amounts of information. The value of the information from each experiment of this kind will progressively increase as more is learned about the functions of each gene and as additional experiments define the global changes in gene expression in diverse other natural processes and genetic perturbations. Perhaps the greatest challenge now is to develop efficient methods for organizing, distributing, interpreting, and extracting insights from the large volumes of data these experiments will provide.

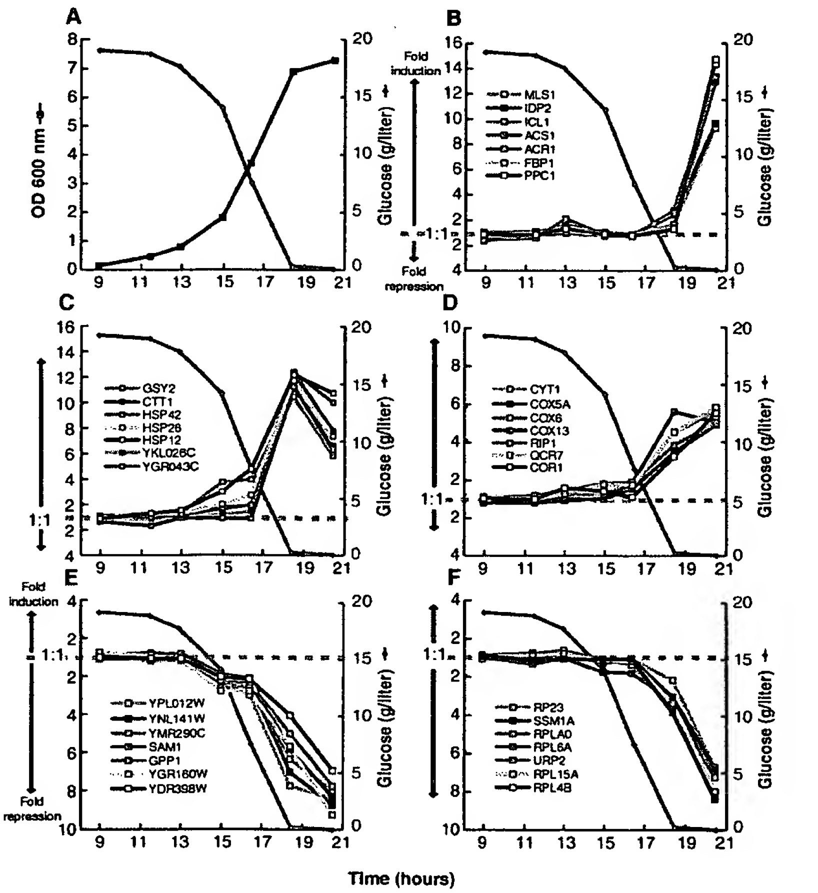


Fig. 5. Distinct temporal patterns of induction or repression help to group genes that share regulatory properties. **(A)** Temporal profile of the cell density, as measured by OD at 600 nm and glucose concentration in the media. **(B)** Seven genes exhibited a strong induction (greater than ninefold) only at the last timepoint (20.5 hours). With the exception of *IDP2*, each of these genes has a CSRE UAS. There were no additional genes observed to match this profile. **(C)** Seven members of a class of genes marked by early induction with a peak in mRNA levels at 18.5 hours. Each of these genes contain STRE motif repeats in their upstream promoter regions. **(D)** Cytochrome c oxidase and ubiquinol cytochrome c reductase genes. Marked by an induction coincident with the diauxic shift, each of these genes contains a consensus binding motif for the HAP2,3,4 protein complex. At least 17 genes shared a similar expression profile. **(E)** *SAM1*, *GPP1*, and several genes of unknown function are repressed before the diauxic shift, and continue to be repressed upon entry into stationary phase. **(F)** Ribosomal protein genes comprise a large class of genes that are repressed upon depletion of glucose. Each of the genes profiled here contains one or more RAP1-binding motifs upstream of its promoter. RAP1 is a transcriptional regulator of most ribosomal proteins.

REFERENCES AND NOTES

1. M. Schena, D. Shalon, R. W. Davis, P. O. Brown, *Science* **270**, 467 (1995).
2. D. Shalon, S. J. Smith, P. O. Brown, *Genome Res.* **6**, 639 (1996).
3. D. Lashkari, *Proc. Natl. Acad. Sci. U.S.A.*, in press.
4. J. DeRisi *et al.*, *Nature Genet.* **14**, 457 (1996).
5. D. J. Lockhart *et al.*, *Nature Biotechnol.* **14**, 1675 (1996).
6. M. Chee *et al.*, *Science* **274**, 610 (1996).
7. M. Johnston and M. Carlson, in *The Molecular Biology of the Yeast* Saccharomyces: *Gene Expression*, E. W. Jones, J. R. Pringle, J. R. Broach, Eds. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1992), p. 193.
8. Primers for each known or predicted protein coding sequence were supplied by Research Genetics. PCR was performed with the protocol supplied by Research Genetics, using genomic DNA from yeast strain S288C as a template. Each PCR product was verified by agarose gel electrophoresis and was deemed correct if the lane contained a single band of appropriate mobility. Failures were marked as such in the database. The overall success rate for a single-pass amplification of 6116 ORFs was ~94.5%.
9. Glass slides (Gold Seal) were cleaned for 2 hours in a solution of 2 N NaOH and 70% ethanol. After rinsing in distilled water, the slides were then treated with a 1:5 dilution of poly-L-lysine adhesive solution (Sigma) for 1 hour, and then dried for 5 min at 40°C in a vacuum oven. DNA samples from 100-µl PCR reactions were purified by ethanol purification in 96-well microtiter plates. The resulting precipitates were resuspended in 3× standard saline citrate (SSC) and transferred to new plates for arraying. A custom-built arraying robot was used to print on a batch of 110 slides. Details of the design of the microarrayer are available at cmgm.stanford.edu/pbrown. After printing, the microarrays were rehydrated for 30 s in a humid chamber and then snap-dried for 2 s on a hot plate (100°C). The DNA was then ultraviolet (UV)-crosslinked to the surface by subjecting the slides to 60 mJ of energy (Stratagene Stratalinker). The rest of the poly-L-lysine surface was blocked by a 15-min incubation in a solution of 70 mM succinic anhydride dissolved in a solution consisting of 315 ml of 1-methyl-2-pyrrolidinone (Aldrich) and 35 ml of 1 M boric acid (pH 8.0). Directly after the blocking reac-

tion, the bound DNA was denatured by a 2-min incubation in distilled water at ~95°C. The slides were then transferred into a bath of 100% ethanol at room temperature, rinsed, and then spun dry in a clinical centrifuge. Slides were stored in a closed box at room temperature until used.
10. YPD medium (8 liters), in a 10-liter fermentation vessel, was inoculated with 2 ml of a fresh overnight culture of yeast strain DBY7286 (MATa, ura3, GAL2). The fermentor was maintained at 30°C with constant agitation and aeration. The glucose content of the media was measured with a UV test kit (Boehringer Mannheim, catalog number 716251) Cell density was measured by OD at 600-nm wavelength. Aliquots of culture were rapidly withdrawn from the fermentation vessel by peristaltic pump, spun down at room temperature, and then flash frozen with liquid nitrogen. Frozen cells were stored at −80°C.
11. Cy3-dUTP or Cy5-dUTP (Amersham) was incorporated during reverse transcription of 1.25 μg of polyadenylated [poly(A)$^+$] RNA, primed by a dT(16) oligomer. This mixture was heated to 70°C for 10 min, and then transferred to ice. A premixed solution, consisting of 200 U Superscript II (Gibco), buffer, deoxyribonucleoside triphosphates, and fluorescent nucleotides, was added to the RNA. Nucleotides were used at these final concentrations: 500 μM for dATP, dCTP, and dGTP and 200 μM for dTTP. Cy3-dUTP and Cy5-dUTP were used at a final concentration of 100 μM. The reaction was then incubated at 42°C for 2 hours. Unincorporated fluorescent nucleotides were removed by first diluting the reaction mixture with of 470 μl of 10 mM tris-HCl (pH 8.0)/1 mM EDTA and then subsequently concentrating the mix to ~5 μl, using Centricon-30 microconcentrators (Amicon).
12. Purified, labeled cDNA was resuspended in 11 μl of 3.5× SSC containing 10 μg poly(dA) and 0.3 μl of 10% SDS. Before hybridization, the solution was boiled for 2 min and then allowed to cool to room temperature. The solution was applied to the microarray under a cover slip, and the slide was placed in a custom hybridization chamber which was subsequently incubated for ~8 to 12 hours in a water bath at 62°C. Before scanning, slides were washed in 2× SSC, 0.2% SDS for 5 min, and then 0.05× SSC for 1 min. Slides were dried before scanning by centrifugation at 500 rpm in a Beckman CS-6R centrifuge.
13. The complete data set is available on the Internet at cmgm.stanford.edu/pbrown/explore/index.html
14. For 95% of all the genes analyzed, the mRNA levels measured in cells harvested at the first and second interval after inoculation differed by a factor of less than 1.5. The correlation coefficient for the comparison between mRNA levels measured for each gene in these two different mRNA samples was 0.98. When duplicate mRNA preparations from the same cell sample were compared in the same way, the correlation coefficient between the expression levels measured for the two samples by comparative hybridization was 0.99.
15. The numbers and identities of known and putative genes, and their homologies to other genes, were gathered from the following public databases: *Saccharomyces* Genome Database (genome-www.stanford.edu), Yeast Protein Database (quest7.proteome.com), and Munich Information Centre for Protein Sequences (speedy.mips.biochem.mpg.de/mips/yeast/index.htmlx).
16. A. Scholer and H. J. Schuller, *Mol. Cell. Biol.* **14**, 3613 (1994).
17. S. Kratzer and H. J. Schuller, *Gene* **161**, 75 (1995).
18. R. J. Haselbeck and H. L. McAlister, *J. Biol. Chem.* **268**, 12116 (1993).
19. M. Fernandez, E. Fernandez, R. Rodicio, *Mol. Gen. Genet.* 242, 727 (1994).
20. A. Hartig *et al.*, *Nucleic Acids Res.* **20**, 5677 (1992).
21. P. M. Martinez *et al.*, *EMBO J.* **15**, 2227 (1996).
22. J. C. Varela, U. M. Praekelt, P. A. Meacock, R. J. Planta, W. H. Mager, *Mol. Cell. Biol.* **15**, 6232 (1995).
23. H. Ruis and C. Schuller, *Bioessays* **17**, 959 (1995).
24. J. L. Parrou, M. A. Teste, J. Francois, *Microbiology* **143**, 1891 (1997).
25. This expression profile was defined as having an induction of greater than 10-fold at 18.5 hours and less than 11-fold at 20.5 hours.
26. S. L. Forsburg and L. Guarente, *Genes Dev.* **3**, 1166 (1989).
27. J. T. Olesen and L. Guarente, *ibid.* **4**, 1714 (1990).
28. M. Rosenkrantz, C. S. Kell, E. A. Pennell, L. J. Devenish, *Mol. Microbiol.* **13**, 119 (1994).
29. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr. The nucleotide codes are as follows: B–C, G, or T; N–G, A, T, or C; R–A or G; and Y–C or T.
30. C. Fondrat and A. Kalogeropoulos, *Comput. Appl. Biosci.* **12**, 363 (1996).
31. D. Shore, *Trends Genet.* **10**, 408 (1994).
32. R. J. Planta and H. A. Raue, *ibid.* **4**, 64 (1988).
33. The degenerate consensus sequence VYCYRNNCMNH was used to search for potential RAP1-binding sites. The exact consensus, as defined by (*30*), is WACAYCCRTACATYW, with up to three differences allowed.
34. S. F. Neuman, S. Bhattacharya, J. R. Broach, *Mol. Cell. Biol.* **15**, 3187 (1995).
35. P. Lesage, X. Yang, M. Carlson, *ibid.* **16**, 1921 (1996).
36. For example, we observed large inductions of the genes coding for *PCK1*, *FBP1* [Z. Yin *et al.*, *Mol. Microbiol.* **20**, 751 (1996)], the central glyoxylate cycle gene *ICL1* [A. Scholer and H. J. Schuller, *Curr. Genet.* **23**, 375 (1993)], and the "aerobic" isoform of acetyl-CoA synthase, *ACS1* [M. A. van den Berg *et al.*, *J. Biol. Chem.* **271**, 28953 (1996)], with concomitant down-regulation of the glycolytic-specific genes *PYK1* and *PFK2* [P. A. Moore *et al.*, *Mol. Cell. Biol.* **11**, 5330 (1991)]. Other genes not directly involved in carbon metabolism but known to be induced upon nutrient limitation include genes encoding cytosolic catalase T *CTT1* [P. H. Bissinger *et al.*, *ibid.* **9**, 1309 (1989)] and several genes encoding small heat-shock proteins, such as *HSP12*, *HSP26*, and *HSP42* [I. Farkas *et al.*, *J. Biol. Chem.* **266**, 15602 (1991); U. M. Praekelt and P. A. Meacock, *Mol. Gen. Genet.* **223**, 97 (1990); D. Wotton *et al.*, *J. Biol. Chem.* **271**, 2717 (1996)].
37. The levels of induction we measured for genes that were expressed at very low levels in the uninduced state (notably, *FBP1* and *PCK1*) were generally lower than those previously reported. This discrepancy was likely due to the conservative background subtraction method we used, which generally resulted in overestimation of very low expression levels (*46*).
38. Cross-hybridization of highly related sequences can also occasionally obscure changes in gene expression, an important concern where members of gene families are functionally specialized and differentially regulated. The major alcohol dehydrogenase genes, *ADH1* and *ADH2*, share 88% nucleotide identity. Reciprocal regulation of these genes is an important feature of the diauxic shift, but was not observed in this experiment, presumably because of cross-hybridization of the fluorescent cDNAs representing these two genes. Nevertheless, we were able to detect differential expression of closely related isoforms of other enzymes, such as *HXK1/HXK2* (77% identical) [P. Herrero *et al.*, *Yeast* **11**, 137 (1995)], *MLS1/DAL7* (73% identical) (*20*), and *PGM1/PGM2* (72% identical) [D. Oh, J. E. Hopper, *Mol. Cell. Biol.* **10**, 1415 (1990)], in accord with previous studies. Use in the microarray of deliberately selected DNA sequences corresponding to the most divergent segments of homologous genes, in lieu of the complete gene sequences, should relieve this problem in many cases.
39. F. E. Williams, U. Varanasi, R. J. Trumbly, *Mol. Cell. Biol.* **11**, 3307 (1991).
40. D. Tzamarias and K. Struhl, *Nature* 369, 758 (1994).
41. Differences in mRNA levels between the tup1Δ and wild-type strain were measured in two independent experiments. The correlation coefficient between the complete sets of expression ratios measured in these duplicate experiments was 0.83. The concordance between the sets of genes that appeared to be induced was very high between the two experiments. When only the 355 genes that showed at least a twofold increase in mRNA in the tup1Δ strain in either of the duplicate experiments were compared, the correlation coefficient was 0.82.
42. The tup1Δ mutation consists of an insertion of the LEU2 coding sequence, including a stop codon, between the ATG of *TUP1* and an Eco RI site 124 base pairs before the stop codon of the *TUP1* gene.
43. L. R. Kowalski, K. Kondo, M. Inouye, *Mol. Microbiol.* **15**, 341 (1995).
44. M. Viswanathan, G. Muthukumar, Y. S. Cong, J. Lenard, *Gene* **148**, 149 (1994).
45. D. Hirata, K. Yano, T. Miyakawa, *Mol. Gen. Genet.* **242**, 250 (1994).
46. A. Gutierrez, L. Caramelo, A. Prieto, M. J. Martinez, A. T. Martinez, *Appl. Environ. Microbiol.* **60**, 1783 (1994).
47. A. Muheim *et al.*, *Eur. J. Biochem.* **195**, 369 (1991).
48. J. A. Wemmie, M. S. Szczypka, D. J. Thiele, W. S. Moye-Rowley, *J. Biol. Chem.* **269**, 32592 (1994).
49. Microarrays were scanned using a custom-built scanning laser microscope built by S. Smith with software written by N. Ziv. Details concerning scanner design and construction are available at cmgm.stanford.edu/pbrown. Images were scanned at a resolution of 20 μm per pixel. A separate scan, using the appropriate excitation line, was done for each of the two fluorophores used. During the scanning process, the ratio between the signals in the two channels was calculated for several array elements containing total genomic DNA. To normalize the two channels with respect to overall intensity, we then adjusted photomultiplier and laser power settings such that the signal ratio at these elements was as close to 1.0 as possible. The combined images were analyzed with custom-written software. A bounding box, fitted to the size of the DNA spots in each quadrant, was placed over each array element. The average fluorescent intensity was calculated by summing the intensities of each pixel present in a bounding box, and then dividing by the total number of pixels. Local area background was calculated for each array element by determining the average fluorescent intensity for the lower 20% of pixel intensities. Although this method tends to underestimate the background, causing an underestimation of extreme ratios, it produces a very consistent and noise-tolerant approximation. Although the analog-to-digital board used for data collection possesses a wide dynamic range (12 bits), several signals were saturated (greater than the maximum signal intensity allowed) at the chosen settings. Therefore, extreme ratios at bright elements are generally underestimated. A signal was deemed significant if the average intensity after background subtraction was at least 2.5-fold higher than the standard deviation in the background measurements for all elements on the array.
50. In addition to the 17 genes shown in Table 1, three additional genes were induced by an average of more than threefold in the duplicate experiments, but in one of the two experiments, the induction was less than twofold (range 1.6- to 1.9-fold)
51. We thank H. Bennett, P. Spellman, J. Ravetto, M. Eisen, R. Pillai, B. Dunn, T. Ferea, and other members of the Brown lab for their assistance and helpful advice. We also thank S. Friend, D. Botstein, S. Smith, J. Hudson, and D. Dolginow for advice, support, and encouragement; K. Struhl and S. Chatterjee for the Tup1 deletion strain; L. Fernandes for helpful advice on Yap1; and S. Klapholz and the reviewers for many helpful comments on the manuscript. Supported by a grant from the National Human Genome Research Institute (NHGRI) (HG00450), and by the Howard Hughes Medical Institute (HHMI). J.D.R. was supported by the HHMI and the NHGRI. V.R. was supported in part by an Institutional Training Grant in Genome Science (T32 HG00044) from the NHGRI. P.O.B. is an associate investigator of the HHMI.

5 September 1997; accepted 22 September 1997

The New York Times
ON THE WEB

October 2, 2003, Thursday

BUSINESS/FINANCIAL DESK

Human Genome Placed on Chip; Biotech Rivals Put It Up for Sale

By ANDREW POLLACK (NYT) 1030 words

The genome on a chip has arrived.

Melding high technology with biology, several companies are rushing to sell slivers of glass or nylon, some as small as postage stamps, packed with pieces of all 30,000 or so known human genes.

The new products will allow scientists to scan all genes in a human tissue sample at once, to determine which genes are active, a job that previously required two or more chips. The whole-genome chips will lower the cost and increase the speed of a widely used test that has transformed biomedical research in the last few years.

"It's sort of a milestone event, very similar to generating an integrated circuit of the genome," said Stephen P. A. Fodor, the chief executive of Affymetrix Inc., the leading seller of gene chips, which are also called microarrays.

Affymetrix, based in Santa Clara, Calif., is expected to announce today that it is accepting orders for its whole-genome chip.

The announcement seems timed to steal some thunder from the rival Agilent Technologies, which is based in nearby Palo Alto. Agilent is to be the host of an analyst meeting today and it plans to announce then that it has started shipping test versions of its whole-genome chip.

Applied Biosystems of Foster City, Calif., a unit of the Applera Corporation, started the race in July with an announcement that it would have a whole-genome chip out by the end of this year. NimbleGen Systems, a small company in Madison, Wis., announced a few days later that it had a genome on a chip that it was not selling but that it was using to run tests for customers.

Gene chips, which detect genes that are active, meaning they are being used to make a protein, have become essential tools. Scientists try to understand the genetic mechanisms of disease by seeing which genes are turned on in, say, a sick kidney or lung compared with those active in a healthy organ. Pharmaceutical companies look at gene activity patterns to try to predict the effects of drugs.

Scientists have found that tumors that look the same under the microscope can differ in terms of which genes are active. So studying gene patterns could become a way to discriminate between deadly and not-so-deadly tumors, or to predict which drug will work best for a particular patient.

Still, even some vendors conceded that the change from two chips to one is more symbolic than revolutionary.

"You can do just as good science with two chips, it costs you a little more," said Roland Green, the vice president for research and development at NimbleGen.

Some scientists questioned whether the chips really have all human genes, because the exact number and identities of all the genes is not known.

The advent of the genome on a chip is, however, evidence that biotechnology, to the extent that it uses electronics, is experiencing some of the rapid progress that has made semiconductors and computers continuously cheaper and smaller.

"One of the effects everyone is looking for in the genomics area is Moore's law -- more data, less money," said Doug Dolginow, an executive vice president at Gene Logic, which sells data from gene chip studies to pharmaceutical companies. "This is a step in that direction."

Moore's law states that the number of transistors on a semiconductor chip doubles every 18 months.

Affymetrix's gene chips are, in fact, made with the same techniques used to make semiconductor chips. In the mid-1990's, the company came out with a set of five chips covering what was then known of the human genome. After the human genome sequence was virtually completed in 2000, the company developed a two-chip set with all the known genes. Now it has the single chip, which some scientists say will be more convenient.

"We like to be able to look at all genes at one time to get a global view of what's going on," said John R. Walker, who runs gene chip operations at the Genomics Institute of the Novartis Research Foundation in San Diego.

Costs should also be lower. Gene chips have been so expensive that many academic scientists still make their own rather than buy them. Affymetrix said it would sell its whole-genome chips for $300 to $500 each, depending on volume, little more than half the price of the two-chip set. The other companies have not announced prices.

For Affymetrix, a successful whole-genome chip "is essential for them to maintain their dominance" of high-end microarrays, said Edward A. Tenthoff, an analyst at U.S. Bancorp Piper Jaffray. Affymetrix had total product sales in 2002 of about $250 million, and a company spokesman said that human genome chips are its top-selling product.

Mr. Tenthoff, who recommends Affymetrix stock, said the company's sales growth rate had moderated as it faces tougher competition. Agilent, a spinoff of Hewlett-Packard that makes its gene chips by printing DNA components onto glass slides using ink jet printers, has gained share, he said. Applied Biosystems, the largest maker of genomics equipment over all, will be

entering the microarray segment of the business with its whole-genome chip, emphasizing the connection of that product to the others it offers, including the gene database developed by its sister company, Celera Genomics.

Jeffrey Trent, scientific director of the Translational Genomics Research Institute in Phoenix, said that while whole-genome chips are useful for medical discovery, the biggest growth of the market will be for chips that can be used by doctors to do diagnoses. And whole-genome chips are too cumbersome for that, he said. Rather, once scientists use the whole-genome chips to find particular genes that are associated with, say, tumor aggressiveness or drug effectiveness, he said, they will then make smaller and cheaper chips containing just those genes for use in diagnosis.

Agilent Technologies

News@Agilent

Agilent Technologies ships whole human genome on single microarray to gene expression customers for evaluation

Company to introduce first commercial whole human microarray by end of year

PALO ALTO, Calif., Oct. 2, 2003

Agilent Technologies Inc. (NYSE: A) today announced it has shipped whole human-genome microarrays to customers for testing and evaluation. The whole genome microarray is based on Agilent's new double-density format, which can accommodate 44,000 features on a single 1" x 3" glass-slide microarray. The new platform enables drug-discovery and disease researchers to perform whole-genome screening at a lower cost and with higher reproducibility.

"This is an important step toward our release of the first whole human-genome microarray product, which is expected to be available for order before the end of the year," said Barney Saunders, vice president and general manager of Agilent's BioResearch Solutions Unit. " Customers have long wanted a one-sample, one-chip format with the increased sensitivity associated with 60-mer probes. The cost savings and high-quality performance make this product a compelling alternative for scientists who make their own microarrays."

Agilent's microarrays are based on the industry-standard 1" x 3" (25mm x 75mm) format, which is compatible with most commercial microarray scanners. All Agilent commercial microarrays are developed using content from public databases and proprietary sources, with full sequence and annotation information made available to customers. Gene sequences for probes are developed using algorithms and then validated empirically through iterative wet-lab testing procedures. The result is a microarray comprised of functionally validated probes, with the most up-to-date and comprehensive genome information commercially available.

Advantages of the double-density format include:

- Lower cost. Not only is one microarray less expensive than two, it requires fewer reagents and reduces instrumentation demands.
- Streamlined workflow. Researchers need prepare and process only one microarray instead of two. This also results in fewer steps in the subsequent data analysis.
- Greater reproducibility. Use of a single microarray further reduces unnecessary variability in experimental conditions.
- Smaller sample use. A smaller quantity of sample material is required to perform an experiment.

Availability

Agilent's Whole Human Genome Microarray is expected to be available for order by the end of the year.

About Agilent Technologies

Agilent Technologies Inc. (NYSE: A) is a global technology leader in communications, electronics, life sciences and chemical analysis. The company's 30,000 employees serve customers in more than 110 countries. Agilent had net revenue of $6 billion in fiscal year 2002. Information about Agilent is available

on the Web at www.agilent.com.

Forward-Looking Statements

This news release contains forward-looking statements (including, without limitation, statements relating to Agilent's expectation that its whole-genome microarray platform will be available for order before the end of 2003) that involve risks and uncertainties that could cause results to differ materially from management's current expectations. These and other risks are detailed in the company's filings with the Securities and Exchange Commission, including its Annual Report on Form 10-K for the year ended Oct. 31, 2002, its Quarterly Report on Form 10-Q for the quarter ended July 31, 2003 and its Current Report on Form 8-K filed Aug. 18, 2003. The company assumes no obligation to update the information in this press release.

###

Contact:

Christina Maehr
+1 408 553 7205
christina_maehr@agilent.com

Today's News

Affymetrix Announces Commercial Launch of Single Array for Human Genome Expression Analysis

AFFYMETRIX GENECHIP(R) BRAND HUMAN GENOME U133 PLUS 2.0 ARRAY

Affymetrix GeneChip(R) Brand Human Genome U133 Plus 2.0 Array.
(PRNewsFoto)[AS]
SANTA CLARA, CA USA 10/02/2003

Website

More Than 1 Million Probes Analyze Expression Levels of Nearly 50,000 RNA Transcripts and Variants on a Single Array the Size of a Thumbnail

SANTA CLARA, Calif., Oct. 2 /PRNewswire/ -- Affymetrix, Inc., (Nasdaq: **AFFX**) announced today that it is taking orders for its new GeneChip(R) brand Human Genome U133 Plus 2.0 Array, offering researchers the protein-coding content of the human genome on a single commercially available catalog microarray. The HG-U133 Plus 2.0 Array analyzes the expression level of nearly 50,000 RNA transcripts and variants with 22 different probes per transcript, providing superior data quality unmatched by technologies using a single probe per transcript.

(Photo: **http://www.newscom.com/cgi-bin/prnh/20031002/SFTH021**)

"With about 1.3 million probes on a chip the size of a human thumbnail, the Human Plus Array represents a leap in array technology data capacity, and further demonstrates the unique power and potential of our technology to explore vast areas of the genome," said Trevor J. Nicholls, Ph.D., Chief Commercial Officer. "Multiple independent measurements for each transcript ensure that our data quality remains the industry standard, even as our data capacity increases dramatically."

The HG-U133 Plus 2.0 Array, which will ship in October, combines the content of the previous HG-U133 two-array set with nearly 10,000 new probe sets representing about 6,500 new genes, for a total of nearly 50,000 RNA transcripts and variants. This new information, verified against the latest version of the publicly available genome map, provides researchers the most comprehensive and up-to-date genome-wide gene expression analysis. The probe design strategy of the HG-U133 Plus 2.0 Array is identical to the previous HG-U133 Set, providing very strong data concordance between the two products. With more than double the data capacity of the previous-generation Affymetrix human product, the HG-U133 Plus 2.0 Array can significantly cut processing and analysis time for scientists in the lab, freeing up valuable resources and accelerating research.

The HG-U133 Plus 2.0 Array sets a new standard for the number of genes and transcripts on any commercially available single array for human gene

expression analysis, while maintaining Affymetrix' unrivaled data quality. The HG-U133 Plus 2.0 Array uses 22 independent measures to detect the hybridization of each transcript on the array, 1.3 million data points in all, more than 30 times that of any other microarray technology. Using multiple, independent measurements provides optimal sensitivity and specificity, and the most accurate, consistent and statistically significant results possible.

"More data points produce more reliable results and ultimately, enable better science," said Nicholls. "Our powerful probe set strategy gives our customers the assurance that their array results actually reflect what's in their sample."

Affymetrix is also launching an updated 11-micron version of its popular 18-micron HG-U133A Array called the GeneChip HG-U133A 2.0 Array. The reduced feature size on this new design means researchers can use smaller sample volumes than on the previous 18-micron array without compromising performance. This new array represents over 20,000 transcripts that can be used to explore human biology and disease processes. All probe sets represented on the original GeneChip HG-U133A Array are identically replicated on the GeneChip HG-U133A 2.0 Array.

More information on the design of the HG-U133 Plus 2.0 Array and the HG-U133A 2.0 Array may be found on the Affymetrix website at **http://www.affymetrix.com**.

Affymetrix will be presenting further information on this and other products at the BioTechnica trade show in Hanover, Germany on Oct. 7-9, 2003. The Company will also hold a press conference on Oct. 7, from 11 a.m. to 12 p.m. at the show regarding the new Human Genome U133 Plus 2.0 Array. If you would like to attend this press conference, please contact Caroline Stupnicka at **c.stupnicka@northbankcommunications.com**.

About Affymetrix:

Affymetrix is a pioneer in creating breakthrough tools that are driving the genomic revolution. By applying the principles of semiconductor technology to the life sciences, Affymetrix develops and commercializes systems that enable scientists to improve the quality of life. The Company's customers include pharmaceutical, biotechnology, agrichemical, diagnostics and consumer products companies as well as academic, government and other non-profit research institutes. Affymetrix offers an expanding portfolio of integrated products and services, including its integrated GeneChip platform, to address growing markets focused on understanding the relationship between genes and human health. Additional information on Affymetrix can be found at **http://www.affymetrix.com**.

All statements in this press release that are not historical are "forward-looking statements" within the meaning of Section 21E of the Securities Exchange Act as amended, including statements regarding Affymetrix' "expectations," "beliefs," "hopes," "intentions," "strategies" or the like. Such statements are subject to risks and uncertainties that could cause actual results to differ materially for Affymetrix from those projected, including, but not limited to risks of the Company's ability to achieve and sustain higher levels of revenue, higher gross margins, reduced operating expenses, uncertainties relating to technological approaches, manufacturing, product development, market acceptance (including uncertainties relating to product development and market acceptance of the GeneChip HG-U133 Human Plus 2.0 Array and the HG-U133A 2.0), personnel retention, uncertainties related to cost and pricing of Affymetrix products, dependence on collaborative partners, uncertainties relating to sole source suppliers, uncertainties relating to FDA and other regulatory approvals, competition, risks relating to intellectual property of others and the uncertainties of patent protection and litigation. These and other risk factors are discussed in Affymetrix' Form 10-K for the

year ended December 31, 2002 and other SEC reports, including its Quarterly Reports on Form 10-Q for subsequent quarterly periods. Affymetrix expressly disclaims any obligation or undertaking to release publicly any updates or revisions to any forward-looking statements contained herein to reflect any change in Affymetrix' expectations with regard thereto or any change in events, conditions, or circumstances on which any such statements are based.

NOTE: Affymetrix, the Affymetrix logo, and GeneChip and are registered trademarks owned or used by Affymetrix, Inc.

SOURCE Affymetrix, Inc.
Web Site: ***http://www.affymetrix.com***
Photo Notes: NewsCom:
http://www.newscom.com/cgi-bin/prnh/20031002/SFTH021 *AP Archive:*
http://photoarchive.ap.org *PRN Photo Desk,*
photodesk@prnewswire.com

)T Vol. 7, No. 15 August 2002

Macroresults through Microarrays

John C. Rockett, Reproductive Toxicology Division (MD-72), National Health and Environmental Effects Research Laboratory, Office of Research and Development, US Environmental Protection Agency, Research Triangle Park, 2525 East Highway 54, Durham, NC 27711, USA; tel: +1 919 541 2071, fax: +1 919 541 4017, e-mail: rockett.john@epa.gov

The third enactment of Cambridge Healthtech Institute's *Macroresults through Microarrays* meeting was held in Boston (MA, USA) from 29 April–1 May 2002. The subtheme of this year's meeting was 'advancing drug discovery', a widely touted application for array technology.

The evolution of microarrays

If you were asked 'Who first conceived of the idea of microarrays', who would come to mind? Mark Schena perhaps, first author of the seminal 1995 paper on cDNA arrays [1]? Maybe Pat Brown, Schena's then supervisor? Or perhaps Stephen Fodor, the primary driver behind Affymetrix's (http://www.affymetrix.com) oligonucleotide-based platform [2]. Brits might even chant the name of Ed Southern [3]. Well, according to Roger Ekins (University College London Medical School; http://www.ucl.ac.uk/medicine/) all these answers would be wrong. It was in fact Ekins and his colleagues who first conceived of and patented 'a new generation of ultrasensitive, miniaturized assays for protein and DNA–RNA measurement based on the use of microarrays' in the mid 1980s [4]. The concept and potential of array technology was more fully described in a later publication, in which Ekins *et al.* [5] concluded that antibody microspots of ~50 μm^2 could be achieved, and that as many as 2 million different immunoassays could, in principle, be accommodated on a surface area of 1 cm^2.

Technological innovation

In practice, it took a different biological molecule (DNA), a different research group, and a leap into microfabrication technology to even begin approaching these kinds of densities [Affymetrix patent 6045996 talks of one million spots cm^{-2}]. Of course, advancing technology is one of the driving engines behind the genomics juggernaut, and we are already seeing '4th generation' machines for fabricating DNA chips. If the company representatives at this meeting are to be believed (and their cases seemed strong), spotting is out, and *in situ* fabrication of oligonucleotide-based 'iterative custom arrays' is in. Whether you go with the Combimatrix's (http://www.combimatrix.com) electrochemically directed synthesis and detection system, febit's (http://www.febit.com) Geniom® technology, or Nimblegen's (http://www.nimblegen.com) Maskless Array Synthesizer technology is a matter of personal choice. However, each of these machines provides the flexibility to design variable length oligonucleotide probes from sequences inputted by the user, and then perform *in situ* synthesis of an array. Each system also boasts unique advantages. For example, Combimatrix's biological array processor is a semiconductor coated with a 3D layer of porous material in which DNA, RNA, peptides or small molecules can be synthesized or immobilized within discrete test sites, while febit's Geniom One® is a fully integrated gene-expression analysis system with minimal user hands-on time – the probe sequences are programmed, the RNA samples inserted, and the gene expression data is pumped out a few hours later.

Cell- and tissue-based arrays

Array technology is in most people's minds firmly linked with gene-expression profiling. Fewer are aware that cell- and tissue-based arrays have been developed, and how they can provide a vital extra dimension to research. In support of this, Barry Bochner gave an update on the cell-based array system that Biolog (http://www.biolog.com) has produced for simultaneously measuring the effects of one gene in the cell under thousands of growth conditions (see [6] for further details). David Walt (Tufts University; http://www.tufts.edu/) is developing single live cell arrays using optical imaging fiber (OIF) technology. An array of microwells is fabricated on the face of an OIF at densities of up to 10 million wells cm^{-2}. Cells are then added to the wells and disperse at an average of one cell per well. Physiological and genetic responses of each cell are measured via fluorescence produced by reporter genes (e.g. *lacZ*, *gfp*. Assays performed so far include yeast live or dead cell assay, microenvironment pH and O_2 measurements, promotor responses using the *lacZ* and *phoA* reporter genes, and protein–protein interactions using the yeast two-hybrid system. The main advantage of this system is that the cells remain alive during the assay, which means a real-time timecourse can be performed and/or the array passed from sample to sample. This would be useful in, for example, the scanning of a combinatorial drug library for specific physiological effects.

Tissue arrays are a useful complementary technology to DNA arrays because they can be used to help validate and

 1359-6446/02/$ – see front matter PII: S1359-6446(02)02352-8

understand the biological and medical significance of gene changes discovered using standard DNA arrays. For example, an array of tumor tissues can be screened for the protein (using immunohistochemistry), message (using *in situ* hybridization) and copy number (using comparative genomic hybridization) of a gene of interest, to determine if expression of the gene (or lack thereof) is related in any way to survival. They can also be used to predict the probability of clinical failure of lead compounds as a result of toxicity by evaluating the distribution of the drug targets in normal tissue. Spyro Mousses and his co-workers at the National Human Genome Research Institute (http://www.nhgri.nih.gov/index.html) have built such arrays, including a multi-tumor array (~5000 specimens, and sections from 36 normal and 800 metastatic tissues) and a normal tissue array (76 tissue and 332 cell types).

The problem with proteins

It has been said that genomics tells us what might happen, transcriptomics indicates what should happen, and proteomics shows what is happening. The impact of functional proteomics on pharmaceutical R&D is rapidly increasing, and protein arrays are being used increasingly in both basic and applied research. Their use lies not only in comparative protein expression and interaction profiling, but also in diagnostics and drug discovery. However, an increasing number of researchers have found that protein arrays, like their cousins the DNA arrays, present several practical obstacles relating to their production and use. For example, in using *Escherichia coli* to produce recombinant eukaryotic proteins from a single expression vector, multiple protein products are often produced, suggesting mixes of truncated or otherwise altered proteins. There is also the obvious concern that the proteins might not be modified in a similar manner to eukaryotic systems. Also, an optimal method for depositing and binding proteins to the selected substrate is yet to be determined, as is the best way to ensure that they are bound in a correctly folded, active conformation.

Several companies have been addressing these problems. Prolinx (http://www.prolinxinc.com) is one such company, and Karin Hughes described their Versalinx™ chemistry for producing protein, peptide and small-molecule arrays. Versalinx™ uses solution-phase conjugation followed by immobilization, resulting in functional orientation of proteins and peptides on the substrate surface. It also offers the valuable additional benefit of exhibiting low non-specific binding. Sense Proteomic (http://www.senseproteomic.com) is also among those addressing these problems to develop robust protein arrays for drug discovery and clinical applications and has developed functional protein array formats based on specific disease tissues. Subtractive hybridization is used to identify genes with altered expression in breast tumor and cystic fibrosis compared to normal tissue. A high throughput cloning strategy (COVET™) is then used to produce libraries of genes that are tagged, cloned, expressed, purified and finally immobilized on glass slides. Initial validation studies have shown that the vast majority of the immobilized proteins do indeed retain biological function.

Stefan Schmidt and his company (GPC Biotech; http://www.gpcbiotech.de) have moved past the platform development stage and, with their focus firmly on drug discovery, are currently developing kinase-profiling arrays. Kinases are important targets for pharmaceutical drug discovery and therapy, and GPC's aim is to simultaneously detect multiple kinases, obtain activity profiles for different cell types, or analyze the ability of drug candidates to inhibit kinase activity. To do this, recombinant kinase substrates are immobilized on membranes, incubated with purified kinase, and the substrates measured for the degree of phosphorylation.

Summary

Meetings like this, packed with exciting discoveries and intriguing and interesting innovation, heavily emphasize the pace at which biotechnology is advancing, to the extent that the number of options for genomic and proteomic researchers can become overwhelming. Although data analysis is perhaps the greatest current concern for array users, an increasing challenge will be to determine the approaches and technology that really work, and to do it in a timely manner.

References

1 Schena, M. *et al.* (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270, 467–470
2 Fodor, S.P. *et al.* (1991) Light-directed, spatially addressable parallel chemical synthesis. *Science* 251, 767–773
3 Southern, E.M. *et al.* (1992) Analyzing and comparing nucleic acid sequences by hybridization to arrays of oligonucleotides: evaluation using experimental models. *Genomics* 13, 1008–1017
4 Ekins, R.P. (1987) US Patent Application 8 803 000
5 Ekins, R. *et al.* (1989) High specific activity chemiluminescent and fluorescent markers: their potential application to high sensitivity and 'multi-analyte' immunoassays. *J. Biolum. Chemilum.* 4, 59–78
6 Rockett, J.C. (2002) Chip, chip, array! Three chips for post-genomic research. *Drug Discov. Today* 7, 458–459

Acknowledgements

I would like to thank Mary Ann Brown (Cambridge Healthtech Institute) and David Dix (US EPA) for critical review of this manuscript prior to submission. This document has been reviewed in accordance with US Environmental Protection Agency policy and approved for publication. Mention of companies, trade names or products does not signify endorsement of such by the EPA.

Electrophoresis 1991, 12, 907–930

N. Leigh Anderson
Ricardo Esquer-Blasco
Jean-Paul Hofmann
Norman G. Anderson

Large Scale Biology Corporation,
Rockville, MD

A two-dimensional gel database of rat liver proteins useful in gene regulation and drug effects studies

A standard two-dimensional (2-D) protein map of Fischer 344 rat liver (F344MST3) is presented, with a tabular listing of more than 1200 protein species. Sodium dodecyl sulfate (SDS) molecular mass and isoelectric point have been established, based on positions of numerous internal standards. This map has been used to connect and compare hundreds of 2-D gels of rat liver samples from a variety of studies, and forms the nucleus of an expanding database describing rat liver proteins and their regulation by various drugs and toxic agents. An example of such a study, involving regulation of cholesterol synthesis by cholesterol-lowering drugs and a high-cholesterol diet, is presented. Since the map has been obtained with a widely used and highly reproducible 2-D gel system (the Iso-Dalt® system), it can be directly related to an expanding body of work in other laboratories.

Contents

1 Introduction

High-resolution two-dimensional electrophoresis of proteins, introduced in 1975 by O'Farrell and others [1–4], has been used over the ensuing 16 years to examine a wide variety of biological systems, the results appearing in more than 5000 published papers. With the advent of computerized systems for analyzing two-dimensional (2-D) gel images and constructing spot databases, it is also possible to plan and assemble integrated bodies of information describing the appearance and regulation of thousands of protein gene products [5, 6]. Creating such databases involves amassing and organizing quantitative data from thousands of 2-D gels, and requires a substantial commitment in technology and resources.

Given the long-term effort required to develop a protein database, the choice of a biological system takes on considerable importance. While *in vitro* systems are ideal for answering many experimental questions, especially in cancer research and genetics, our experience with cell cultures and tissue samples suggests that some *in vivo* approaches could have major advantages. In particular, we have noticed that liver tissue samples from rats and mice appear to show greater quantitative reproducibility (in terms of individual protein expression) than replicate cell cultures. This is perhaps a natural result of the homeostasis maintained in a complete animal *vs.* the well-known variability of cell cultures, the latter due principally to differences in reagents (*e.g.*, fetal bovine serum), conditions (*e.g.*, pH) and genetic "evolution" of cell lines while in culture. It is also more difficult to generate adequate amounts of protein from cell culture systems (particularly with attached cells), forcing the investigator to resort to radioisotope-based or silver-based stain-detection methods. While these methods are more sensitive (sometimes much more sensitive) than the Coomassie Brilliant Blue (CBB) stain typically used for protein detection in "large" protein samples, they are generally more variable, more labor-intensive and, in the case of radiographic methods, may generate highly "noisy" images, due to the properties of the films used. By contrast, large protein samples can easily be prepared from liver using urea/Nonidet P-40 (NP-40) solubilization and stained with CBB, which has the advantage of being easily reproducible [8]. Finally, there remains the question of the "truthfulness" of many *in vitro* systems as compared to their *in vivo* analogs; how great are the changes caused by the introduction into a cul-

Correspondence: Dr. N. Leigh Anderson, Large Scale Biology Corporation, 9620 Medical Center Drive, Rockville, MD 20850, USA

Abbreviations: CBB, Coomassie Brilliant Blue; CPK, creatine phosphokinase; 2-D, two-dimensional; IEF, isoelectric focusing; MSN, master spot number; NP-40, Nonidet P-40, SDS, sodium dodecyl sulfate

0173-0835/91/1111-0907 $3.50+.25/0

ture and the associated shift to strong selection for growth, and how do these affect experimental outcomes? Hence the apparent advantages of *in vitro* systems, in terms of experimental manipulation, may be counterbalanced by other factors relating to 2-D data quality.

There is a second important class of reasons for exploring the use of an *in vivo* biological system such as the liver. Historically, there have been two broad approaches to the mechanistic dissection of biochemical processes in intact cellular systems: genetics (a search for informative mutants) and the use of chemical agents (drugs and chemical toxins). Both approaches help us to understand complex systems by disrupting some specific functional element and showing us the result. With the development of techniques for genetic manipulation and cloning, the genetic approach can be effectively applied either *in vitro* or *in vivo*, although the *in vitro* route is usually quicker. The chemical approach can also be applied to either sort of biological system; here, however, the bulk of consistently acquired information is in experimental animals (rats and mice). While most biologists know a short list of compounds having specific, experimentally useful effects (*e.g.*, inhibitors of protein synthesis, ionophores, polymerase inhibitors, channel blockers, nucleotide analogs, and compounds affecting polymerization of cytoskeletal proteins), there is a much larger number of interesting chemically-induced effects, most of them characterized by toxicologists and pharmacologists in rodent systems. Just as a thorough genetic analysis would involve saturating a genome with mutations, it is possible to imagine a saturating number of drugs, the analysis of whose actions would reveal the complete biochemistry of the cell. While organized drug discovery efforts usually target specific desired effects, the nature of the process, with its dependence on screening large numbers of compounds, necessarily produces many unanticipated effects. It is therefore reasonable to suppose that the required broad range of compounds necessary to achieve "biochemical saturation" may be forthcoming; in fact, it may already exist among the hundreds of thousands of compounds that failed to qualify as drugs.

Among organs, the liver is an obvious choice for the study of chemical effects because of its well-known plasticity and responsiveness. The brain appears to be quite plastic (*e.g.* [7]), but it is a complicated mixture of cell types requiring skillful dissection for most experiments. The kidney, while quite responsive, also presents a potentially confounding mixture of cell types. The liver, by contrast, is made up of one predominant cell type which is easy to solubilize: the hepatocyte, representing more than 95% of its mass. Most importantly, the liver performs many homeostatic functions that require rapid modulation of gene expression. It appears that most chemical agents tested affect gene expression in the liver at some dosage (N. Leigh Anderson, unpublished observations), an interesting contrast to our earlier work with lymphocytes, for example, which seem to be much less responsive. Such results conform to the expectation that cells with a homeostatic, physiological role should be more plastic than cells differentiated for a purpose dependent on the action of a limited number of specific genes.

The liver also allows the parallels between *in vitro* and *in vivo* systems to be examined in detail. Significant progress has been made in the development of mouse, rat and human hepatocyte culture systems, as well as in precision-cut tissue slices. Using such an array of techniques, it is now possible to assemble a matrix of mammalian systems including mouse and rat *in vivo* on one level and mouse, rat and human *in vitro* on a second level, and to compare effects between species and between systems. This approach allows us to draw informed conclusions regarding the biochemical "universality" of biological responses among the mammals and to offer some insight into the validity of *in vitro* approaches for toxicological screening. We believe this data will be necessary if *in vitro* alternatives are to achieve wide usage in government-mandated safety testing of drugs, consumer products and industrial and agricultural chemicals.

A number of interesting studies have been published using 2-D mapping to examine effects in the rodent liver. A number of investigarors have made use of the technique to screen for existing genetic variants [8–11] or induced mutations [12–14], mainly in the mouse. This work builds on the wealth of genetic information available on the mouse and its established position as a mammalian mutation-detection system. While some studies of chemical effects have been undertaken in the mouse [15–17], most have used the rat [18–23]. The examination of the cytochrome p-450 system, in particular, has been carried out almost exclusively on the rat [24, 25].

These considerations lead us to conclude that rodent liver offers the best opportunity to systematically examine an array of gene regulation systems, and ultimately to build a predictive model of large-scale mammalian gene control. The basic underlying foundation of such a project is a reliable, reproducible master 2-D pattern of liver, to which ongoing experimental results can be referred. In this paper, we report such a master pattern for the acidic and neutral proteins of rat liver (pattern F344MST3). In future, this master will be supplemented by maps of basic proteins, and analogous maps of mouse and human liver.

2 Materials and methods

2.1 Sample preparation

Liver is an ideal sample material for most biochemical studies, including 2-D analysis. A sample is taken of approximately 0.5 g of tissue from the apical end of the left lobe of the liver. Solubilization is effected as rapidly as practical: a delay of 5–15 min appears to cause no major alteration in liver protein composition if the liver pieces are kept cold (*e.g.*, on ice) in the interim. In the solubilization process, the liver sample is weighed, placed in a glass homogenizer (*e.g.*, 15 mL Wheaton); 8 volumes of solubilizing solution*

* The solubilizing solution is composed of 2% NP-40 (Sigma), 9 M urea (analytical grade, *e.g.*, BDH or Bio-Rad), 0.5% dithiothreitol (DTT; Sigma) and 2% carrier ampholytes (pH 9–11 LKB: these come as a 20% stock solution, so 2% final concentration is achieved by making the final solution 10% 9–11 Ampholine by volume). A large batch of solubilizer (several hundred mL) is made and stored frozen at −80 °C in aliquots sufficient to provide enough for one day's estimated sample preparation requirement. The solution is never allowed to become warmer than room temperature at any stage during preparation or thawing for use, since heating of concentrated urea solutions can produce contaminants that covalently modify proteins producing artifactual charge shifts. Once thawed, any unused solubilizer is discarded.

...dded (*i.e.*, 4 mL per 0.5 g tissue) and the mixture is ho-...genized using first the loose- and then then the tight-fit-...g glass pestle. This takes approximately 5 strokes with ...ch pestle and is carried out at room temperature because ...ea would crystallize out in the cold. Once the liver sample ...thoroughly homogenized in the solubilizer, it is assumed ...at all the proteins are denatured (by the chaotropic effect ...the urea and NP-40 detergent) and the enzymes inacti-...ted by the high pH (~9.5). Therefore these samples may ...kept at room temperature until they can be centrifuged ...frozen as a group (within several hours of preparation). ...e samples are centrifuged for 6×10^6 g min (*e.g.*, 500 000 g for 12 min using a Beckman TL-100 centrifuge). The ...ntrifuge rotor is maintained at just below room tempera-...re (*e.g.*, 15–20°C), but not too cold, so as to prevent the ...ecipitation of urea. The centrifuge of choice is a Beckman ...L-100 because of the sample tube sizes available, but any ...tracentrifuge accepting smallish tubes will suffice. When ...appropriate centrifuge is not available near the site of ...mple preparation, samples can be frozen at −80°C and ...awed prior to centrifugation and collection of superna-...nts. Each supernatant is carefully removed following cen-...ifugation and aliquoted into at least 4 clean tubes for stor-...ge. This is done by transferring all the supernatant to one ...lean tube, mixing this gently (to assure homogeneous ...omposition) and then dividing it into 4 aliquots. The ali-...uots are frozen immediately at −80°C. These multiple ali-...uots can provide insurance against a failed run or a freezer ...reakdown.

2.2 Two-dimensional electrophoresis

...ample proteins are resolved by 2-D electrophoresis using ...he 20 × 25 cm Iso-Dalt[3] 2-D gel system ([26–29]; pro-...duced by LSB and by Hoefer Scientific Instruments, San ...Francisco) operating with 20 gels per batch. All first-dimen-...sional isoelectric focusing (IEF) gels are prepared using the ...same single standardized batch of carrier ampholytes ...BDH 4–8A in the present case, selected by LSB's batch-...testing program for rat and mouse database work**). A 10 ...μL sample of solubilized liver protein is applied to each gel, ...and the gels are run for 33 000 to 34 500 volt-hours using a ...progressively increasing voltage protocol implemented by ...a programmable high-voltage power supply. An Ange-...lique™ computer-controlled gradient-casting system (pro-...duced by LSB) is used to prepare second-dimensional sod-...ium dodecyl sulfate (SDS) polyacrylamide gradient slab ...gels in which the top 5% of the gel is 11%T acrylamide, and ...the lower 95% of the gel varies linearly from 11% to 18%T.

...This system has recently been modified so as to employ a ...commercially available 30.8%T acrylamide/*N,N'*-methyle-...nebisacrylamide prepared solution (thus avoiding the han-...dling of the solid acrylamide monomer) and three addi-...tional stock solutions: buffer (made from Sigma pre-set ...Tris), persulfate and *N,N,N',N'*-tetramethylethylenedi-...amine (TEMED). Each gel is identified by a computer-...printed filter paper label polymerized into the lower left cor-...ner of the gel. First-dimensional IEF tube gels are loaded directly (as extruded) onto the slab gels without equilibration, and held in place by polyester fabric wedges (Wedgies™, produced by LSB) to avoid the use of hot agarose. Second-dimensional slab gels are run overnight, in groups of 20, in cooled DALT tanks (10°C) with buffer circulation. All run parameters, reagent source and lot information, and notations of deviation from expected results are entered by the technician responsible on a detailed, multi-page record of the experiment.

** This material (succeeding certified batches of which are available from Hoefer Scientific Instruments) has the most linear pH gradient produced by any ampholyte tested except for the Pharmacia wide range (which has an unacceptable tendency to bind high-molecular weight acidic proteins, causing them to streak).

2.3 Staining

Following SDS-electrophoresis, slab gels are stained for protein using a colloidal Coomassie Blue G-250 procedure in covered plastic boxes, with 10 gels (totalling approximately 1 L of gel) per box. This procedure (based on the work of Neuhoff [30, 31]) involves fixation in 1.5 L of 50% ethanol and 2% phosphoric acid for 2 h, three 30 min washes, each in 2 L of cold tap water, and transfer to 1.5 L of 34% methanol, 17% ammonium sulfate and 2% phosphoric acid for 1 h, followed by the addition of a gram of powdered Coomassie Blue G-250 stain. Staining requires approximately 4 days to reach equilibrium intensity, whereupon gels are transferred to cool tap water and their surfaces rinsed to remove any particulate stain prior to scanning. Gels may be kept for several months in water with added sodium azide. The water washes remove ethanol that would dissolve the stain (and render the system noncolloidal, with high backgrounds). The concentrated ammonium sulfate and methanol solution is diluted by equilibration with the water volume of the gels to automatically achieve the correct final concentrations for colloidal staining. Practical advantages of this staining approach can be summarized as follows: (i) the low, flat background makes computer evaluation of small spots (max OD < 0.02) possible, especially when using laser densitometry; (ii) up to 1500 spots can be reliably detected on many gels (*e.g.*, rat liver) at loadings low enough to preserve excellent resolution; and (iii) reproducibility appears to be very good: at least several hundred spots have coefficients of reproducibility less than 15%. This value is at least as good as previous CBB methods, and significantly better than many silver stain systems.

2.4 Positional standardization

The carbamylated rabbit muscle creatine phosphokinase (CPK) standards [32] are purchased from Pharmacia and BDH. Amino acid compositions, and numbers of residues present in proteins used for internal standardization, are taken from the Protein Identification Resource (PIR) sequence database [33].

2.5 Computer analysis

Stained slab gels are digitized in red light at 134 micron resolution, using either a Molecular Dynamics laser scanner (with pixel sampling) or an Eikonix 78/99 CCD scanner. Raw digitized gel images are archived on high-density DAT tape (or equivalent storage media) and a greyscale videoprint prepared from the raw digital image as hard-copy backup of the gel image. Gels are processed using the Kepler[4] software system (produced by LSB), a commercially available workstation-based software package built on

some of the principles of the earlier TYCHO system [34–41]. Procedure PROC008 is used to yield a spotlist giving position, shape and density information for each detected spot. This procedure makes use of digital filtering, mathematical morphology techniques and digital masking to remove the background, and uses full 2-D least-squares optimization to refine the parameters of a 2-D Gaussian shape for each spot. Processing parameters and file locations are stored in a relational database, while various log files detailing operation of the automatic analysis software are archived with the reduced data. The computed resolution and level of Gaussian convergence of each gel are inspected and archived for quality control purposes.

Experiment packages are constructed using the Kepler experiment definition database to assemble groups of 2-D patterns corresponding to the experimental groups (*e.g.*, treated and control animals). Each 2-D pattern is matched to the appropriate "master" 2-D pattern (pattern F344MST3 in the case of Fischer 344 rat liver), thereby providing linkage to the existing rodent protein 2-D databases. The software allows experiments containing hundreds of gels to be constructed and analyzed as a unit, with up to 100 gels displayed on the screen at one time for comparative purposes and multiple pages to accommodate experiments of > 1000 gels. For each treatment, proteins showing significant quantitative differences *vs.* appropriate controls are selected using group-wise statistical parameters (*e.g.*, Student's t-test, Kepler® procedure STUDENT). Proteins satisfying various quantitative criteria (such as $P <$ 0.001 difference from appropriate controls) are represented as highlighted spots onscreen or on computer-plotted protein maps and stored as spot populations (*i.e.*, logical vectors) in a liver protein database. Quantitative data (spot parameters, statistical or other computed values) are stored as real-valued vectors in the database. Analysis of coregulation is performed using a Pierson product-moment correlation (Kepler procedure CORREL) to determine whether groups of proteins are coordinately regulated by any of the treatments. Such groups can be presented graphically on a protein map, and reported together with the statistical criteria used to assess the level of coregulation. Multivariate statistical analysis (*e.g.*, principal components' analysis) is performed on data exported to SAS (SAS Institute).

2.6 Graphical data output

Graphical results are prepared in GKS and translated within Kepler® into output for any of a variety of devices. Linedrawing output is typically prepared as Postscript and printed on an Apple Laserwriter. Detailed maps presented here have been generated using an ultra-high-resolution Postscript-compatible Linotronic output device. Greyscale graphics are reproduced from the workstation screen using a Seikosha videoprinter. Patterns are shown in the standard orientation, with high molecular mass at the top and acidic proteins to the left.

2.7 Experiment LSBC04

In the study described here 12-week-old Charles River male F344 rats were used. Diets were prepared at LSB, based on a Purina 5755M Basal Purified Diet. Lovastatin and cholestyramine were obtained as prescription pharmaceuticals, ground and mixed with the diet at concentrations of 0.075% and 1%, respectively. The high cholesterol diet was Purina 5801M-A (5% cholesterol plus 1% sodium cholate in the control diet). Animal work was carried out by Microbiological Associates (Bethesda, MD). Animals were acclimatized for one week on the control diet, fed test or control diets for one week, and sacrificed on day 8. Average daily doses of lovastatin and cholestyramine in appropriate groups were 37 mg/kg/day and 5 g/kg/day, respectively, based on the weight of the food consumed. Liver samples were collected and prepared for 2-D electrophoresis according to the standard liver protocol (homogenization in 8 volumes of 9 M urea, 2% NP-40, 0.5% dithiothreitol, 2% LKB pH 9–11 carrier ampholytes, followed by centrifugation for 30 min at 80000 × *g*). Kidney, brain and plasma samples were frozen. Gels were run as described above, and the data was analyzed using the Kepler® system. Gels were scaled, to remove the effect of differences in protein loading, by setting the summed abundances of a large number of matched spots equal for each gel (linear scaling).

3 Results and discussion

3.1 The rat liver protein 2-D map

F344MST3 is a standard 2-D pattern of rat liver proteins, based on the Fischer 344 strain. This pattern was initiated from a single 2-D gel and extensively edited in an experiment comparing it to a range of protein loads, so as to include both small spots and well-resolved representations of high-abundance spots. More than 700 rat liver 2-D patterns have been matched to F344MST3 in a series of drug effects and protein characterization experiments, and numerous new spots (induced by specific drugs, for instance) have been added as a result. A modified version including additional spots present in the Sprague-Dawley outbred rat has also been developed (data not shown). Figure 1 shows a greyscale representation and Fig. 2 a schematic plot of the master pattern. More than 1200 spots are included, most of which are visible on typical gels loaded with 10 µL of solubilized liver protein prepared by the standard method and stained with colloidal Coomassie Blue. Master spot numbers (MSN's) have been assigned to all proteins, and appear in the following figures, each showing one quadrant of the pattern. Figure 3 shows the upper left (acidic, high molecular mass) quadrant, Fig. 4 the upper right (basic, high molecular mass) quadrant, Fig. 5 the lower left (acidic, low molecular mass) quadrant, and Fig. 6 the lower right (basic, low molecular mass) quadrant. The quadrants overlap as an aid to moving between them. The gel position (in 100 micron units), isoelectric point (relative to the CPK internal p*I* standards) and SDS molecular mass (from the calibration curve in Fig. 8) are listed for each spot (Table 1). Because of the precision of the CPK-p*I* values, these parameters can be used to relate spot locations between gel systems more reliably than using p*I* measurements expressed as pH. A major objective of current studies is the identification of all major spots corresponding to known liver proteins, as well as rigorous definitions of subcellular organelle contents. Of particular interest to us is the parallel development of identifications in the rat and mouse liver maps, allowing detailed comparisons of gene expression effects in the two systems. The results of these studies will be presented systematically in a later edition of this database.

we include here a useful series of 22 orienting identifi-tions as an aid to other users of the rat liver pattern (Table ...

2 Carbamylated charge standards, computed p*I*'s and molecular mass standardization

'e have previously shown that the use of a system of close-spaced internal p*I* markers (made by carbamylating a asic protein) offers an accurate and workable solution to ie problem of assigning positions in the p*I* dimension [32]. he same system, based on 36 protein species made by car-amylating rabbit muscle CPK has been used here to as-ign p*I*'s to most rat liver acidic and neutral proteins. The tandards were coelectrophoresed with total liver proteins, nd the standard spots added to a special version of the naster pattern F344MST3. The gel *X*-coordinates of all ver protein spots lying within the CPK charge train were hen transformed into CPK p*I* positions by interpolation etween the positions of immediately adjacent standards Table 1) using a Kepler[2] vector procedure.

t has proven possible to compute fairly accurate p*I* values or many proteins from the amino acid composition [42]. We have attempted here to test a further elaboration of this pproach, in which we computed p*I*'s for the CPK standards hemselves, based on our knowledge of the rabbit muscle CPK sequence and the fact that adjacent members of the harge train typically differ by blockage of one additional ly-sine residue (Table 3). We compared these values to similar computed p*I*'s for an additional set of carbamylated stand-ards made from human hemoglobin beta chains and a se-ries of rat liver and human plasma proteins of known posi-tion and sequence (Fig. 7, Table 4). The result demonstrates good concordance between these systems. Two proteins show significant deviations: liver fatty-acid binding protein (FABP; #1 in Table 4) and protein disulphide isomerase (#20 in the table). The FABP spot present on F344MST3 may represent a charge-modified version of a more basic parent spot closer to the expected p*I*, not resolved in the IEF/SDS gel. Of particular importance is the fact that, by comparing computed p*I*'s of sequenced but unlocated pro-teins with the CPK p*I*'s, we can assign a probable gel loca-tion without making any assumptions regarding the actual gel pH gradient. This offers a useful shortcut, given the va-garies of pH measurement on small diameter IEF gels. We have used this approach to compute the CPK p*I*'s of all rat and mouse proteins in the PIR sequence database, as an aid to protein identification (data not shown).

In order to standardize SDS molecular weight (SDS-MW), we have used a standard curve fitted to a series of identified proteins (Fig. 8). Rather than using molecular mass *per se*, we have elected to use the number of amino acids in the polypeptide chain, as perhaps a better indication of the length of the SDS-coated rod that is sieved by the second dimension slab. The resulting values were multiplied by 112 (the weighted average mass of amino acids in se-quenced proteins) to give predicted molecular masses. Be-cause we use gradient slabs, we have not constrained the fit-ted curve to conform to any predetermined model; rather ... tried many equations and selected the best using the program "Tablecurve" on a PC. The equation chosen was $y = a + bx + c/x^2$, where y is the number of residues, x is the gel *Y* coordinate, a is 511.83, b is -0.2731 and c is 33183801. The resulting fit appears to be fairly good over a broad range of molecular mass.

3.3 An example of rat liver gene regulation: Cholesterol metabolism

Experiment LSBC04 was designed as a small-scale test of the regulation of cholesterol metabolism *in vivo* by three agents included in the diet: lovastatin (Mevacor[3], an inhibi-tor of HMG-CoA reductase); cholestyramine (a bile acid sequestrant that has the effect of removing cholesterol from the gut-liver recirculation); and cholesterol itself. The first two agents should lower available cholesterol and the third should raise it, allowing manipulation of relevant gene expression control systems in both directions. Such an experiment offers an interesting test of the 2-D mapping system since most of the pathway enzymes are present in low abundance, many are membrane-bound and difficult to solubilize, and the pathway itself is complex. Approxima-tely 1000 proteins were separated and detected in liver ho-mogenates. Twenty-one proteins were found to be affected by at least one treatment, and these could be divided into several coregulated groups.

3.3.1 MSN 413 (putative cytosolic HMG-CoA synthase) and sets of spots regulated coordinately or inversely

One group of spots (including a spot assigned to the cyto-solic HMG-CoA synthase, MSN 413) showed the expected increase in abundance with lovastatin or cholestyramine, the synergistic further increase with lovastatin and choles-tyramine, and a dramatic decrease with the high cholesterol diet. Spot number 413 is the most strongly regulated pro-tein in the present experiment, showing a 5- to 10-fold in-duction after a 1 week treatment with 0.075% lovastatin and 1% cholestyramine in the diet (Figs. 9 and 10). Its expres-sion follows precisely the expectation for an enzyme whose abundance is controlled by the cholesterol level; it is pro-gressively increased from the control levels by cholestyra-mine, lovastatin and lovastatin plus cholestyramine, and it sinks below the threshold of detection in animals fed the high cholesterol diet. This spot has been tentatively identi-fied as the cytosolic HMG-CoA synthase, based on a reac-tion with an antiserum to that protein provided by Dr. Mi-chael Greenspan at Merck Sharp & Dohme Research Labo-ratories. This enzyme lies immediately before HMG-CoA reductase in the liver cholesterol biosynthesis pathway, and is known to be co-regulated with it. Spot 413 has an SDS molecular weight of about 54 000 and a CPK p*I* of -11.4, in reasonably close agreement with a molecular weight of 57300 and a CPK p*I* of -15.7 computed from the known se-quence of the hamster enzyme [43].

Using a classical product-moment correlation test (Kepler procedure CORREL), a series of five additional spots was found to be coregulated with 413. The level of correlation was exceedingly high (> 95%). Two of these, 1250 and 933, are at similar molecular weights and approximately one charge more acidic than 413 (Fig. 9), indicating that they may be covalently modified forms of the 413 polypeptide. This suspicion is strengthened by the observation that both spots are also stained by the antibody to cytosolic HMG-CoA synthase. The remaining three correlated spots appear

to comprise an additional related pair (1253 and 1001) of around 40 kDa and a single spot (1119) of around 28 kDa. Because these two presumed proteins are present at substantially lower abundances than 413, and because the cytosolic HMG-CoA synthase is reported to consist of only one type of polypeptide, they are likely to represent other, very tightly coregulated enzymes. A second group of six spots was selected based on a regulatory pattern close to the inverse of that for spot 413 (MSN's 34, 79, 178, 182, 204, 347; data not shown). For these proteins, the lowest level of expression occurs with exposure to lovastatin plus cholestyramine and the highest level upon exposure to the high-cholesterol diet. Spots 182 and 79 are highly correlated and lie about one charge apart at the same molecular weight; they may thus be isoforms of a single protein. The other four spots probably represent additional enzymes or subunits.

3.3.2 MSN 235 and coregulated spots

A third group of five spots, mainly comprised of mitochondrial proteins including putative mitochondrial HMG-CoA synthase spots, showed a modest induction by lovastatin alone, but little or no effect with any of the other treatments (including the combination of lovastatin and cholestyramine; Fig. 12). This result is intriguing because lovastatin was expected to affect only the regulation of enzymes of cholesterol synthesis, which is entirely extra-mitochondrial. Three of the spots (235, 134, 144) form a closely-packed triad at approximately 30 kDa, and are likely to represent isoforms of one protein. All three spots are stained by an antibody to the mitochondrial form of HMG-CoA synthase obtained from Dr. Greenspan. Subcellular fractionation indicates a mitochondrial location. The other two spots (633 at about 38 kDa and 724 at about 69 kDa) are each present at lower abundance than the members of the triad.

3.3.3 An example of an anti-synergistic effect

A sixth spot (367) shows strong induction by lovastatin (two- to threefold), and about half as much induction with lovastatin plus cholestyramine, but without sharing the animal-animal heterogeneity pattern of the 235-set (Fig. 13). This protein is also mitochondrial, and represents the clearest example of an anti-synergistic effect of lovastatin and cholestyramine. The existence of such an effect demonstrates that lovastatin and cholestyramine do not act exclusively through the same regulatory pathway.

3.3.4 Complexity of the cholesterol synthesis pathway

Taken together, these results suggest that treatment with lovastatin alone can affect both cytosolic and mitochondrial pathways using HMG-CoA, while cholestyramine, on the other hand, either alone or in combination with lovastatin, produces a strong effect on the putative cytosolic pathway, but little or no effect on the putative mitochondrial pathway. An explanation for this difference may lie in lovastatin's effect on levels of HMG-CoA and related precursor compounds that are exchanged between the cytosol and the mitochondrion, whereas cholestyramine should affect only the cytosolic pathways directly controlled by cholesterol and bile acid levels. It remains to be explained why some proteins of the putative mitochondrial pathway are so much more variable in their expression in all groups. An examination of all the coregulated groups suggests that quantitative statistical techniques can extract a wealth of interesting information from large sets of reproducible gels. The abundance of spots in the 413 coregulation group, for example, shows an amazing level of concordance in their relative expression among the five individuals of the lovastatin and cholestyramine treatment group. This effect is not due to differences in total protein loading, since they have already been removed by scaling, and since proteins with quite different regulation patterns can be demonstrated (e.g., Fig. 13). Such effects raise the possibility that many gene coregulation sets may be revealed through the study of a sufficiently large population of control animals (*i.e.*, without any experimental manipulation). This approach, exploiting natural biological variation in protein expression instead of drug effects, offers an important incentive for the construction of a large library of control animal patterns.

4 Conclusions

Because of the widespread use of rat liver in both basic biochemistry and in toxicology, there is a long-term need for a comprehensive database of liver proteins. The rat liver master pattern presented here has proven to be an accurate representation of this system, having been matched to more than 700 gels to date. As the number of proteins identified and the number of compounds tested for gene expression effects grows, we expect this database to contribute valuable insights into gene regulation. Its practical utility in several areas of mechanistic toxicology is already being demonstrated.

Received September 11, 1991

5 References

[1] O'Farrell, P., *J. Biol. Chem.* 1975, *250*, 4007–4021.
[2] Klose, J., *Humangenetik* 1975, *26*, 231–243.
[3] Scheele, G. A., *J. Biol. Chem.* 1975, *250*, 5375–5385.
[4] Iborra, G. and Buhler, J. M., *Anal. Biochem.* 1976, *74*, 503–511.
[5] Anderson, N. G. and Anderson, N. L., *Behring. Inst. Mitt.* 1979, *63*, 169–210
[6] Anderson, N. G and Anderson, N. L., *Clin. Chem.* 1982, *28*, 739–748
[7] Heydorn, W. E., Creed, G. J. and Jacobowitz, D. M., *J. Pharmacol. Exp. Therap.* 1984, *229*, 622–628.
[8] Anderson, N. L., Nance, S. L., Tollaksen, S. L., Giere, F. A. and Anderson, N. G., *Electrophoresis* 1985, *6*, 592–599.
[9] Racine, R. R. and Langley, C. H., *Biochem. Genet.* 1980, *18*, 185–197.
[10] Klose, J., *Mol. Evol.* 1982, *18*, 315–328.
[11] Neel, J. V., Baier, L., Hanash, S. and Erickson, R. P., *J. Hered.* 1985, *76*, 314–320.
[12] Marshall, R. R., Raj, A. S., Grant, F. J. and Heddle, J. A., *Can. J. Genet. Cytol.* 1983, *25*, 457–446.
[13] Taylor, J., Anderson, N. L., Anderson, N. G., Gemmell, A., Giometti, C. S., Nance, S. L. and Tollaksen, S. L., in: Dunn, M. J. (Ed.), *Electrophoresis '86*, Verlag Chemie, Weinheim 1986, pp. 583–587.
[14] Giometti, C. S., Gemmell, M. A., Nance, S. L., Tollaksen, S. L. and Taylor, J., *J. Biol. Chem.* 1987, *262*, 12764–12767.
[15] Anderson, N. L., Giere, F. A., Nance, S. L., Gemmell, M. A., Tollaksen, S. L. and Anderson, N. G., in: Galteau, M.-M. and Siest, G. (Eds.), *Progrès Récents en Electrophorèse Bidimensionelle*, Presses Universitaires de Nancy, Nancy 1986, pp. 253–260.
[16] Anderson, N. L., Swanson, M., Giere, F. A., Tollaksen, S., Gemmell, A., Nance, S. L. and Anderson, N. G., *Electrophoresis* 1986, *7*, 44–48.

Anderson, N. L., Giere, F. A., Nance, S. L., Gemmell, M. A., Tollaksen, S. L. and Anderson, N. G., *Fundam. Appl. Toxicol.* 1987, *8*, 39–50.

Anderson, N. L., in: *New Horizons in Toxicology*, Eli Lilly Symposium, 1991, in press.

Antoine, B., Rahimi-Pour, A., Siest, G., Magdalou, J. and Galteau, M. M., *Cell. Biochem. Funct.* 1987, *5*, 217–231.

Elliott, B. M., Ramasamy, R., Stonard, M. D. and Spragg, S. P. *Biochim. Biophys. Acta* 1986, *876*, 135–140.

Huber, B. E., Heilman, C. A., Wirth, P. J., Miller, M. J. and Thorgeirsson, S. S., *Hepatology* 1986, *6*, 209–219.

Wirth, P. J. and Vesterberg, O., *Electrophoresis* 1988, *9*, 47–53.

Witzmann, F. A. and Parker, D. N., *Toxicol. Lett.* 1991, *57*, 29–36.

Rampersaud, A., Waxman, D. J., Ryan, D. E., Levin, W. and Walz, F. G., Jr., *Arch. Biochem. Biophys.* 1985, *243*, 174–183.

Vlasuk, G. P. and Walz, F. G., Jr., *Anal. Biochem.* 1980, *105*, 112–120.

Anderson, N. G. and Anderson, N. L., *Anal. Biochem.* 1978, *85*, 331–340.

Anderson, N. L. and Anderson, N. G., *Anal. Biochem.* 1978, *85*, 341–354.

Anderson, L., Hofmann, J.-P., Anderson, E., Walker, B. and Anderson, N. G., in: Endler, A. T. and Hanash, S. (Eds.), *Two-Dimensional Electrophoresis*, VCH Verlagsgesellschaft, Weinheim 1989, pp. 288–297.

Anderson, L., *Two-Dimensional Electrophoresis: Operation of the ISO-DALT® System*. Large Scale Biology Press, Washington, DC 1988, ISBN 0-945532-00-8, 170pp.

Neuhoff, V., Stamm, R. and Eibl, H., *Electrophoresis* 1985, *6*, 427–448.

[31] Neuhoff, V., Arold, N., Taube, D. and Ehrhardt, W., *Electrophoresis* 1988, *9*, 255–262.

[32] Anderson, N. L. and Hickman, B. J., *Anal. Biochem.* 1979, *93*, 312–320.

[33] Sidman, K. E., George, D. E., Barker, W. C. and Hunt, L. T., *Nucl. Acids Res.* 1988, *16*, 1869–1871.

[34] Taylor, J., Anderson, N. L., Coulter, B. P., Scandora, A. E. and Anderson, N. G., in: Radola, B. J. (Ed.), *Electrophoresis '79*, de Gruyter, Berlin 1980, pp. 329–339.

[35] Taylor, J., Anderson, N. L. and Anderson, N. G., in: Allen, R. C. and Arnaud, P. (Eds.), *Electrophoresis '81*, de Gruyter, Berlin 1981, pp. 383–400.

[36] Anderson, N. L., Taylor, J., Scandora, A. E., Coulter, B. P. and Anderson, N. G., *Clin. Chem.* 1981, *27*, 1807–1820.

[37] Taylor, J., Anderson, N. L., Scandora, A. E., Jr., Willard, K. E. and Anderson, N. G., *Clin. Chem.* 1982, *28*, 861–866.

[38] Taylor, J., Anderson, N. L. and Anderson, N. G., *Electrophoresis* 1983, *4*, 338–345.

[39] Anderson, N. L. and Taylor, J., in: *Proceedings of the Fourth Annual Conference and Exposition of the National Computer Graphics Association*, Chicago, June 26–30, 1983, pp. 69–76.

[40] Anderson, N. L., Hofmann, J.-P., Gemmell, A. and Taylor, J., *Clin. Chem.* 1984, *30*, 2031–2036.

[41] Anderson, L., in: Schafer-Nielsen, C. (Ed.), *Electrophoresis '88*, VCH Verlagsgesellschaft, Weinheim 1988, pp. 313–321.

[42] Neidhardt, F. C., Appleby, D. A., Sankar, P., Hutton, M. E. and Phillips, T. A., *Electrophoresis* 1989, *10*, 116–121.

[43] Gil, G., Goldstein, J. L., Slaughter, C. A. and Brown, M. S., *J. Biol. Chem.* 1986, *261*, 3710–3716.

6 Addendum 1: Figures 1–13

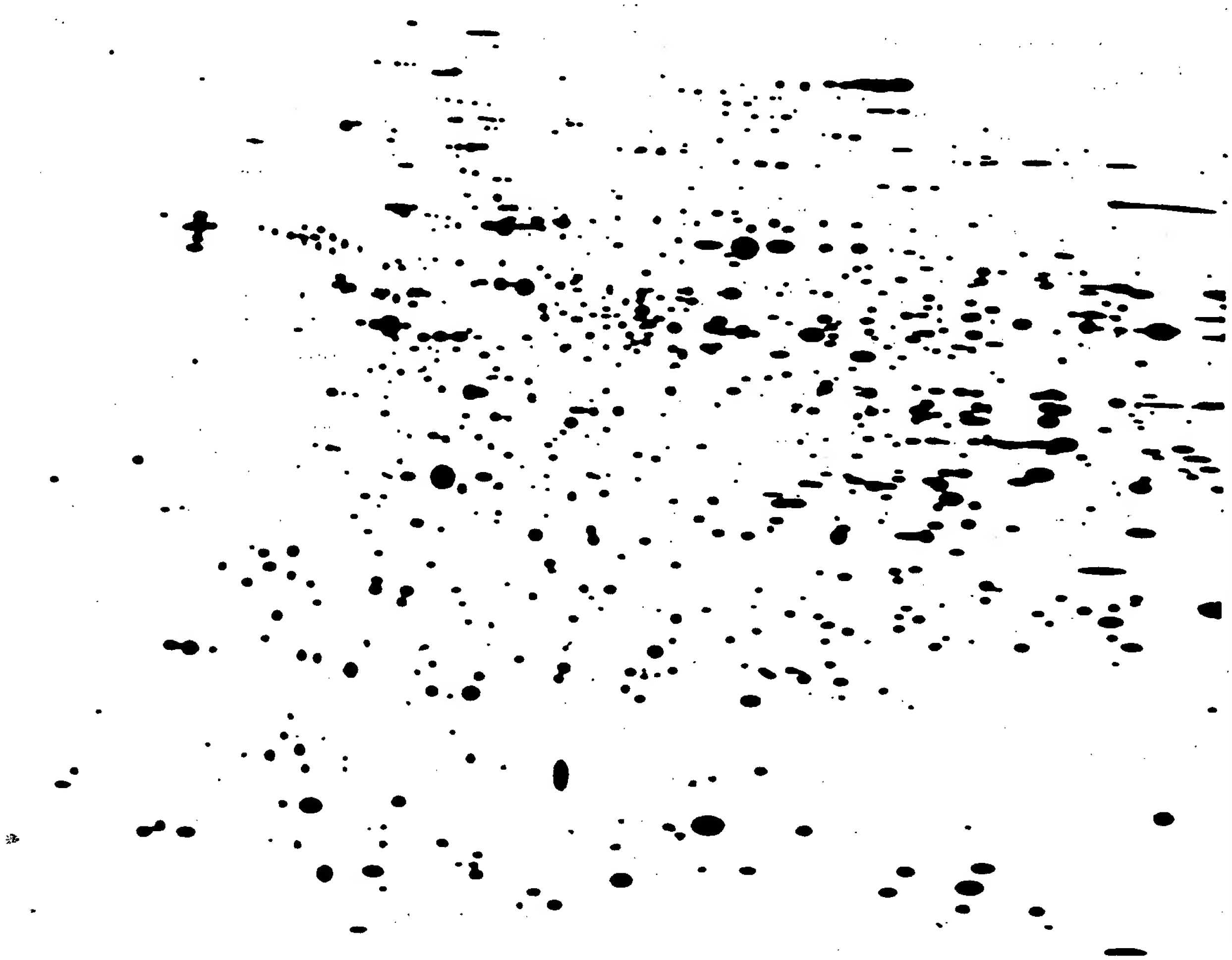
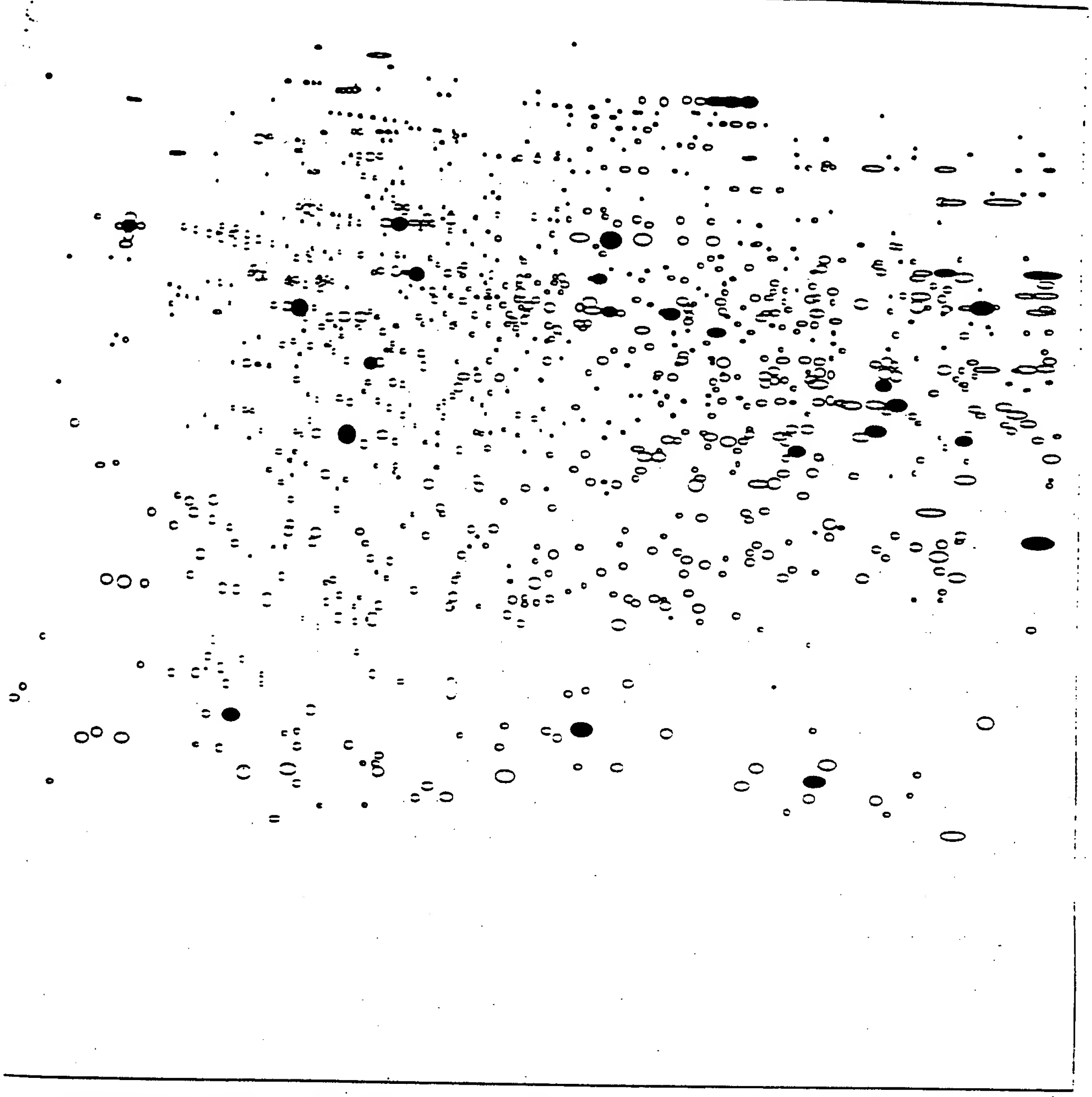


Figure 1. Synthetic representation of the standard rat liver 2-D master pattern, rendered as a greyscale image using a videoprinter.



…re 2. Schematic representation of the master pattern (the same as Fig. 1), useful as an aid in relating specific areas of Fig. 1 and the following detailed …rants.

1

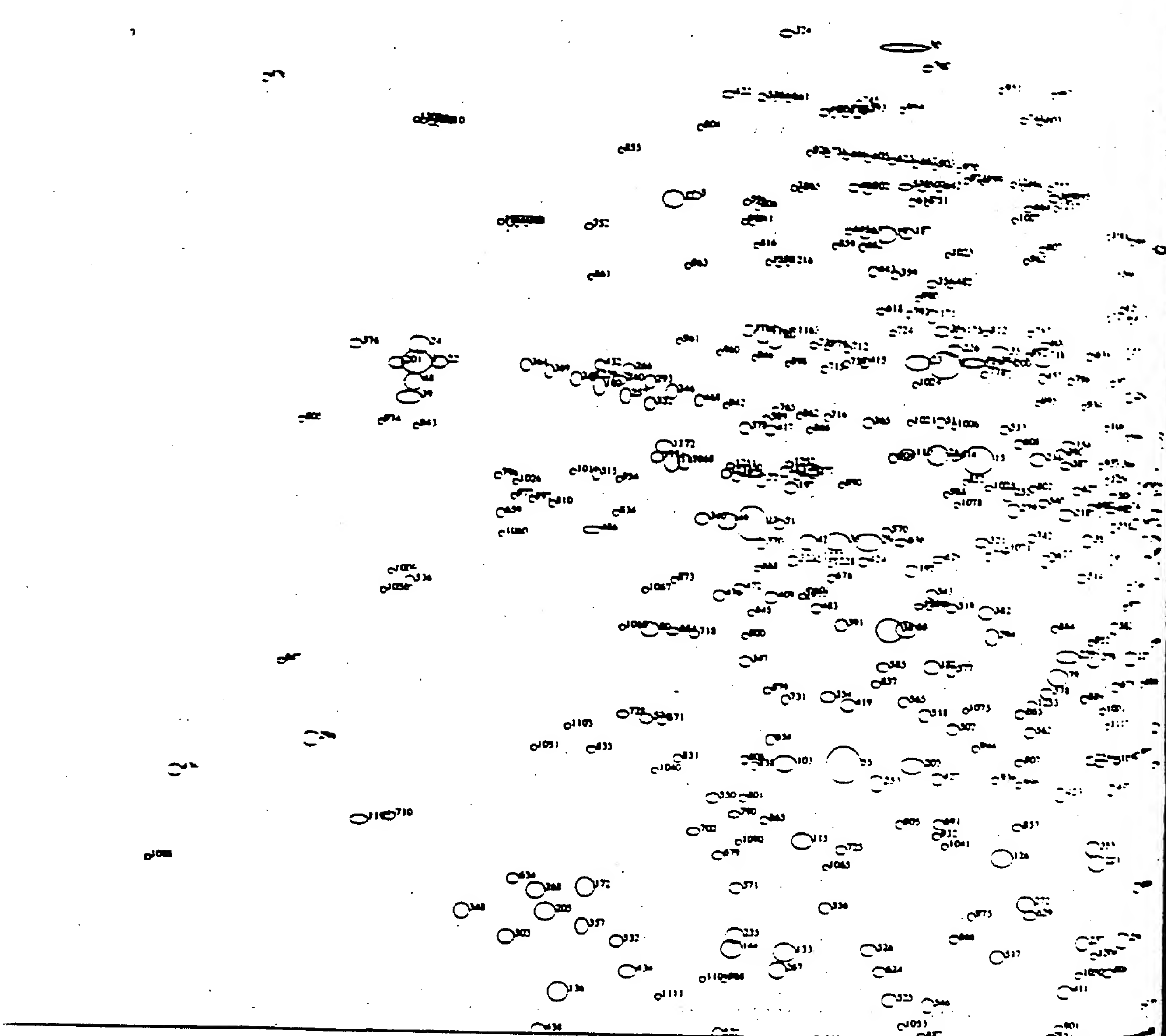
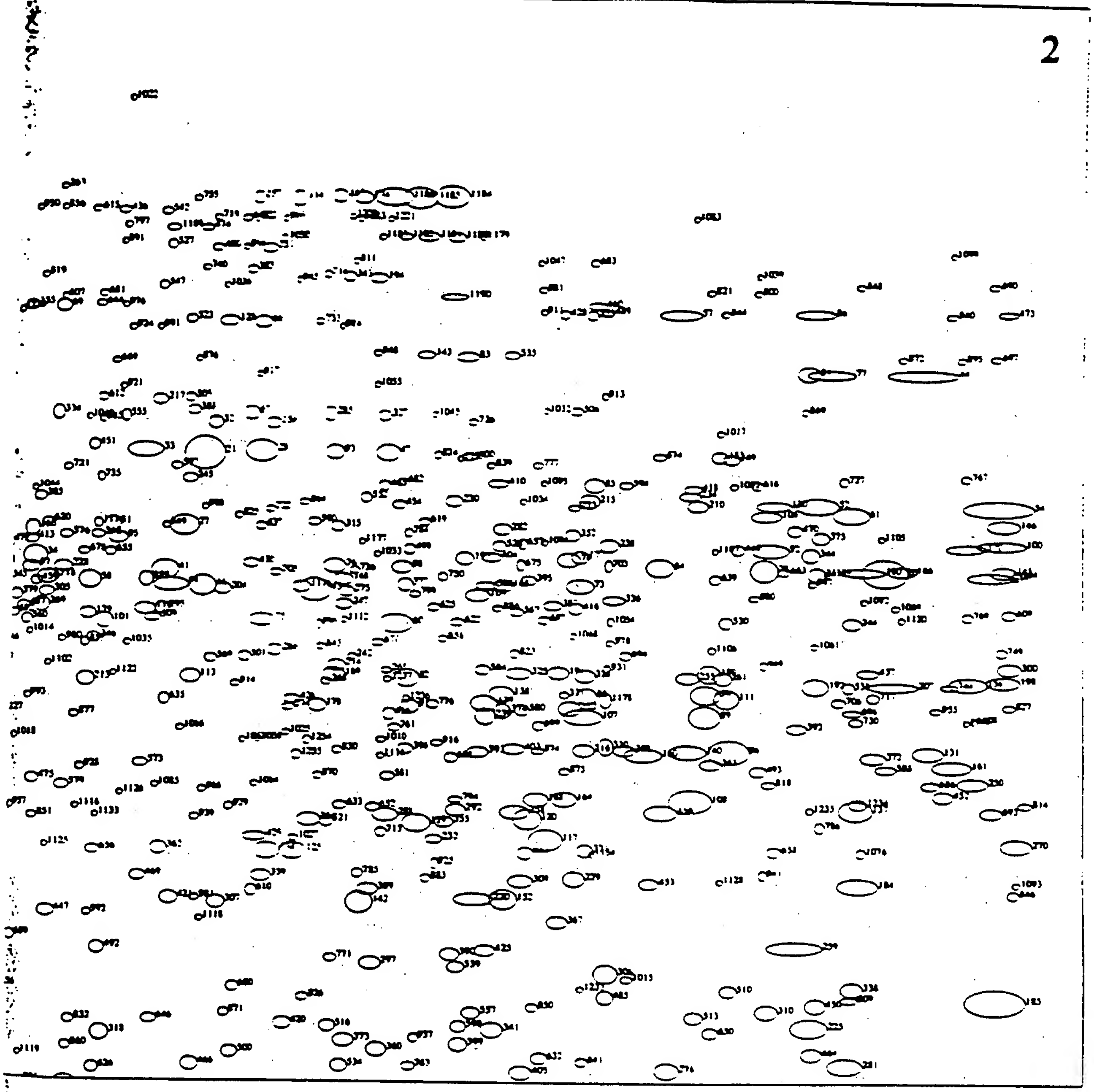


Figure 3. Upper left (high molecular weight, acidic) quadrant (#1) of the rat liver map, showing spot numbers.



...ure 4. Upper right (high molecular weight, basic) quadrant (#2) of the rat liver map, showing spot numbers.

Electrophoresis 1991, 12, 907-9[illegible]

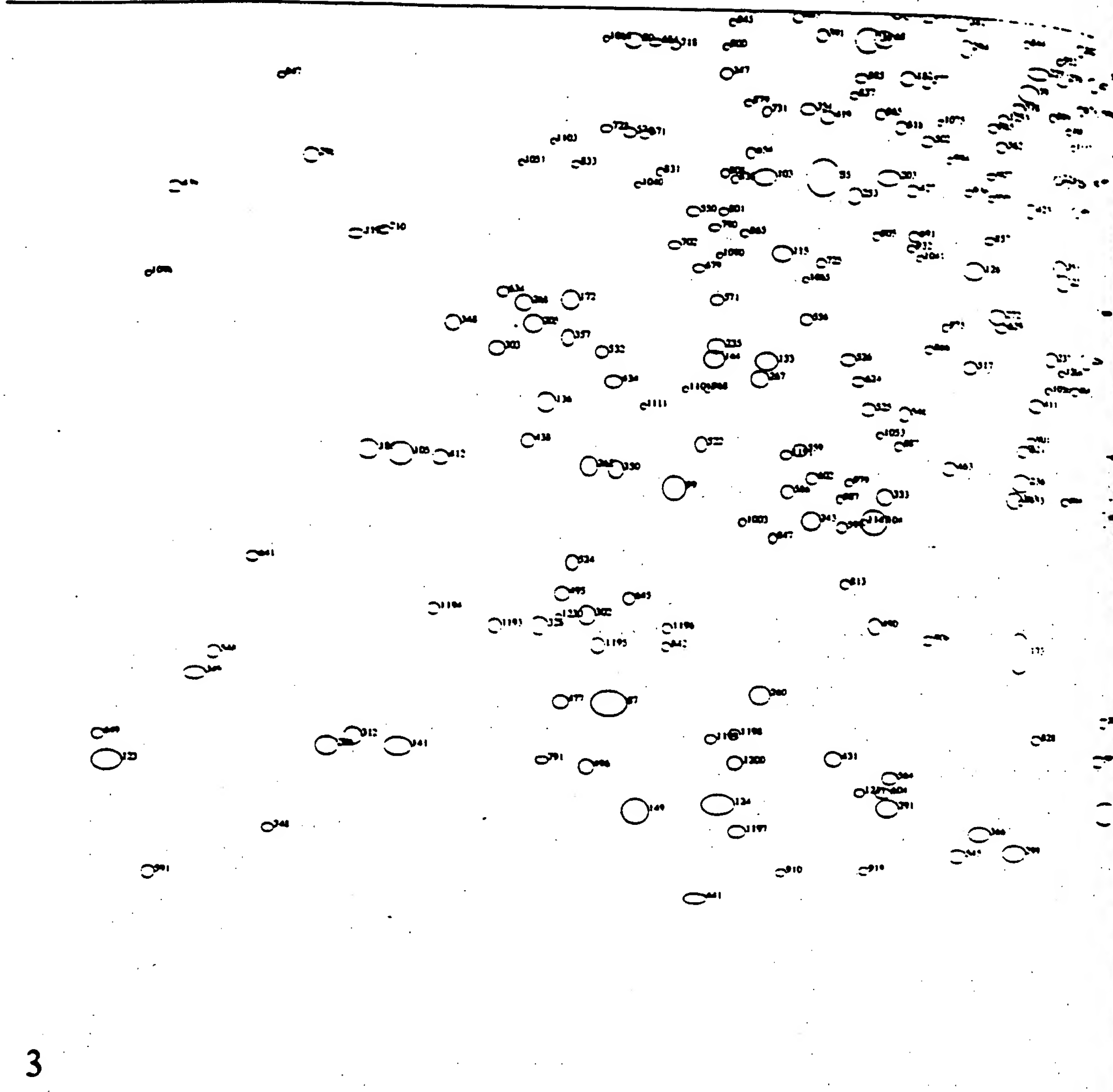
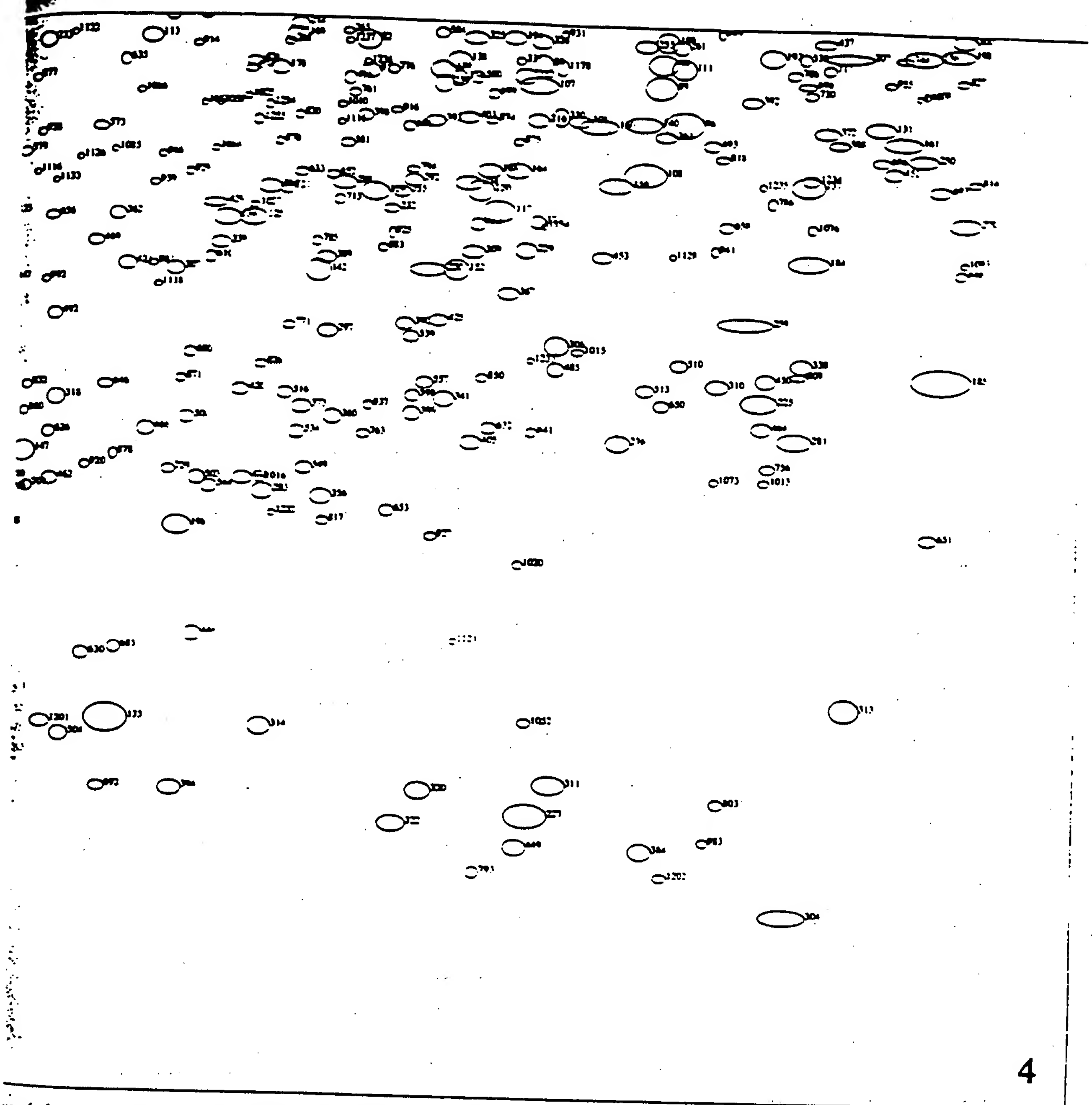
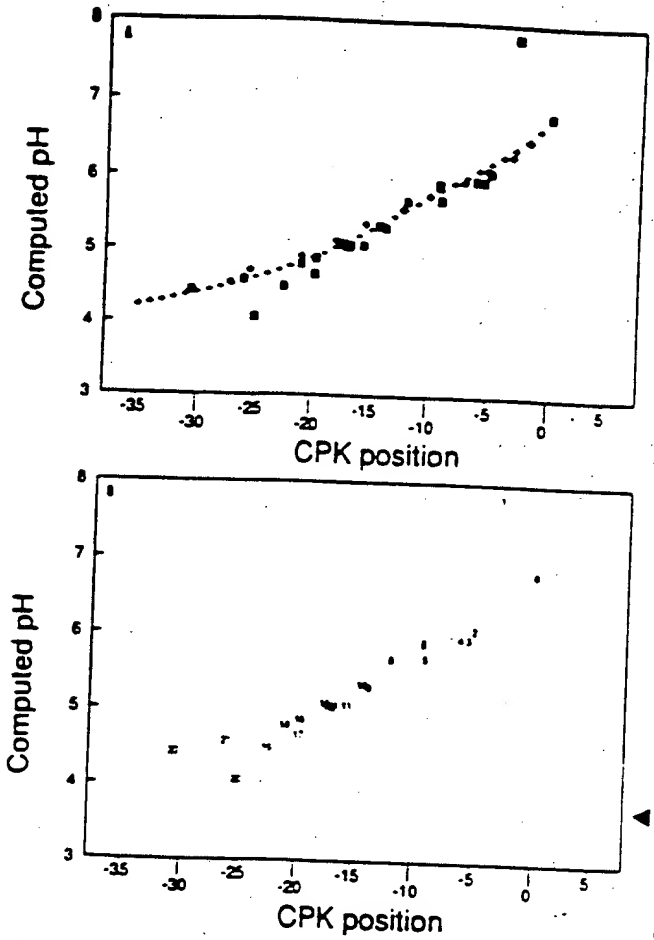


Figure 5. Lower left (low molecular weight, acidic) quadrant (#3) of the rat liver map, showing spot numbers.



...ure 6. Lower right (low molecular weight, basic) quadrant (#4) of the rat liver map, showing spot numbers.



◀ *Figure 7.* (a) Plot of computed isoelectric point versus gel X-position for two sets of carbamylated standard proteins (rabbit muscle CPK [+] and human hemoglobin β chain, filled diamonds) and several other proteins (shaded squares). (b) The identities of the various proteins represented by the squares are indicated by the numbers in corresponding positions on (a); these refer to Table 4.

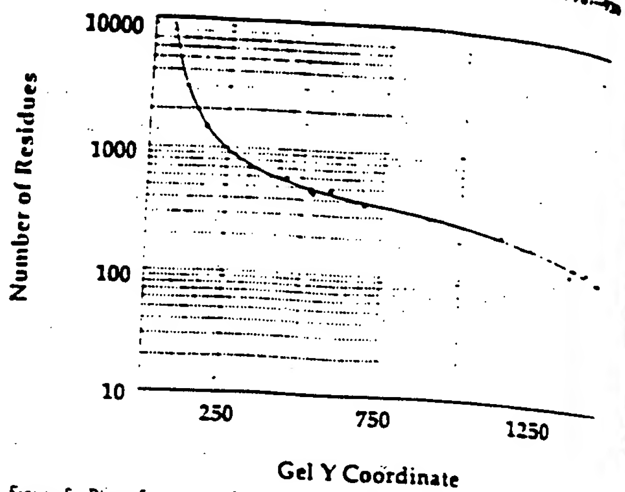


Figure 8. Plot of number of amino acids versus gel Y-position, with fitted curve used to predict molecular mass of unidentified proteins

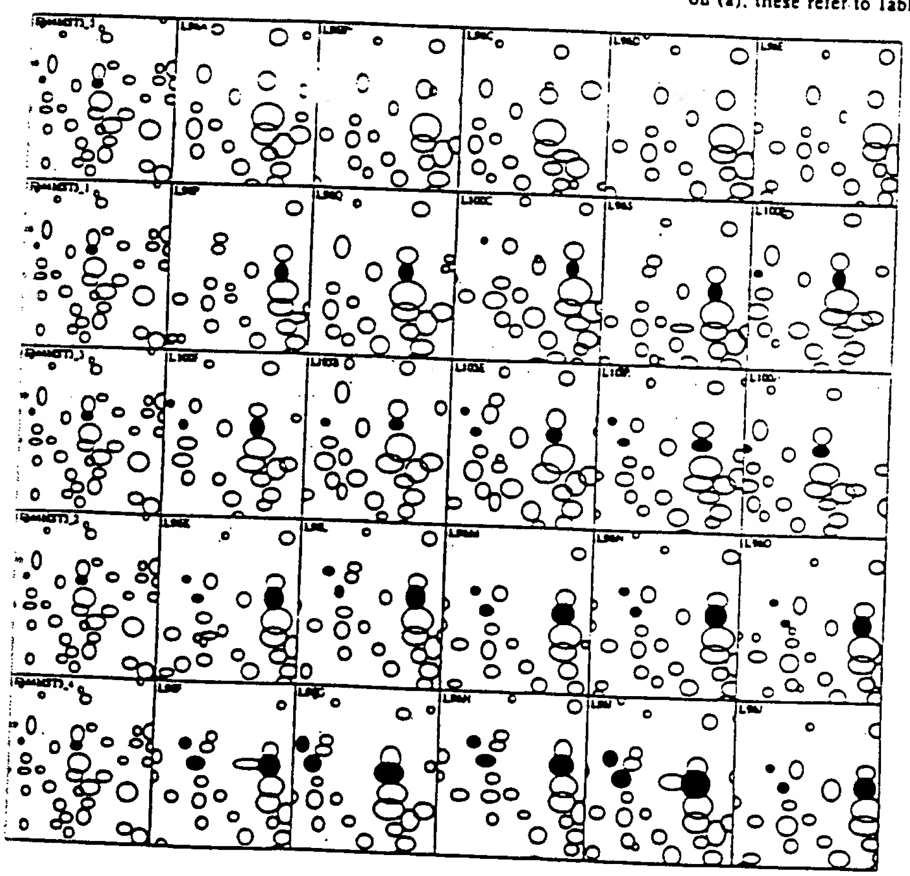


Figure 9. Montage showing effects in the region of MSN:413. The montage shows a small window into one portion of the 2-D pattern, one row of windows for each experimental group, and one panel for each gel in the experiment. The left-most pattern in each row is a group-specific copy of the master pattern followed by the patterns for the five individual rats in the group. The highlighted protein spots (filled circles) are spot 413 (on the right of each panel; identified as cytosolic HMG-CoA synthase) and two modified forms of it (1250 and 933). From the top, the rows (experimental groups) are: high cholesterol, controls, cholestyramine, lovastatin, and lovastatin plus cholestyramine.

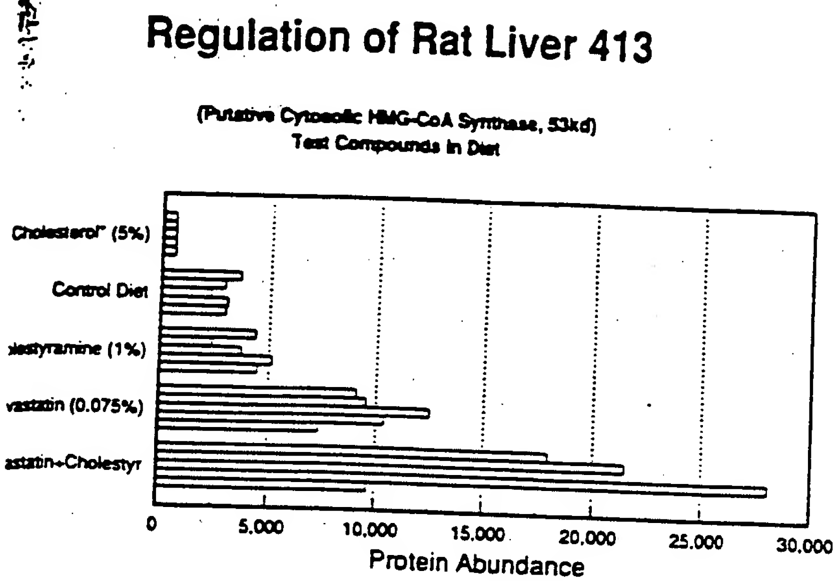


Figure 10. Bargraph showing the quantitative effects of various treatments on the abundance of MSN:413 (cytosolic HMG-CoA synthase) in the gels of Fig. 9.

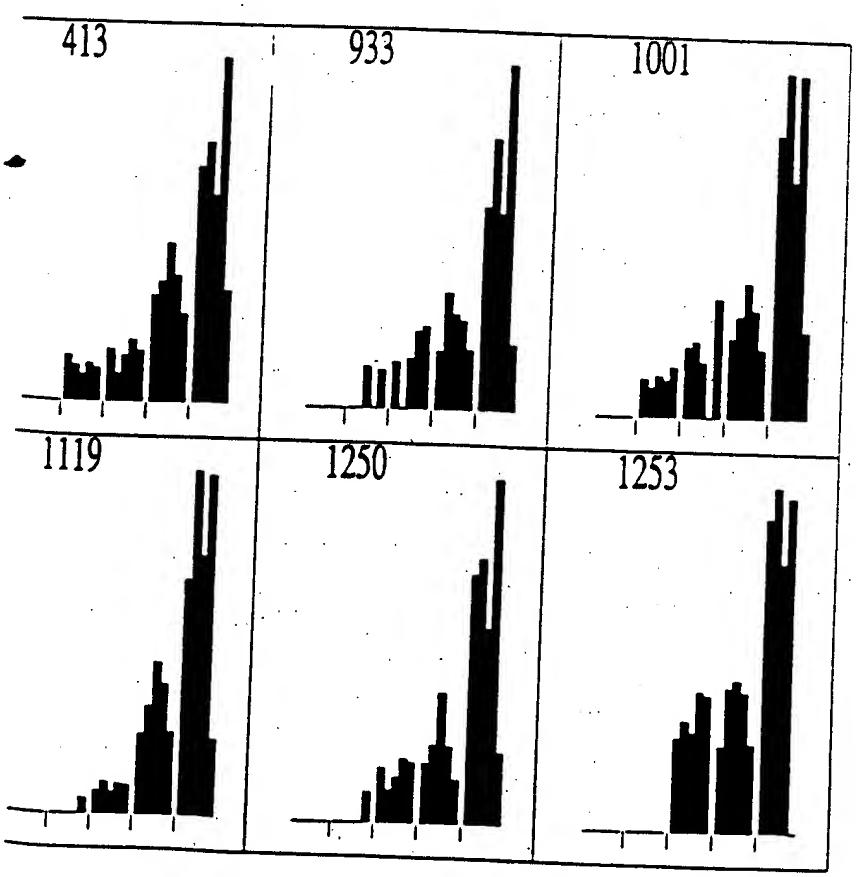


Figure 11. Bargraphs of a series of six coregulated spots including MSN:413. In the bargraphs, the abundances of the appropriate spot (master spot number shown at the top of the panel) in each animal are shown. The five five-animal groups are in the order (left to right): high cholesterol, controls, cholestyramine, lovastatin, and lovastatin plus cholestyramine. Each bar within a group represents one experimental animal liver (one 2-D gel). Note the correlated expression of the 6 spots, especially in the two far right (most strongly induced) groups.

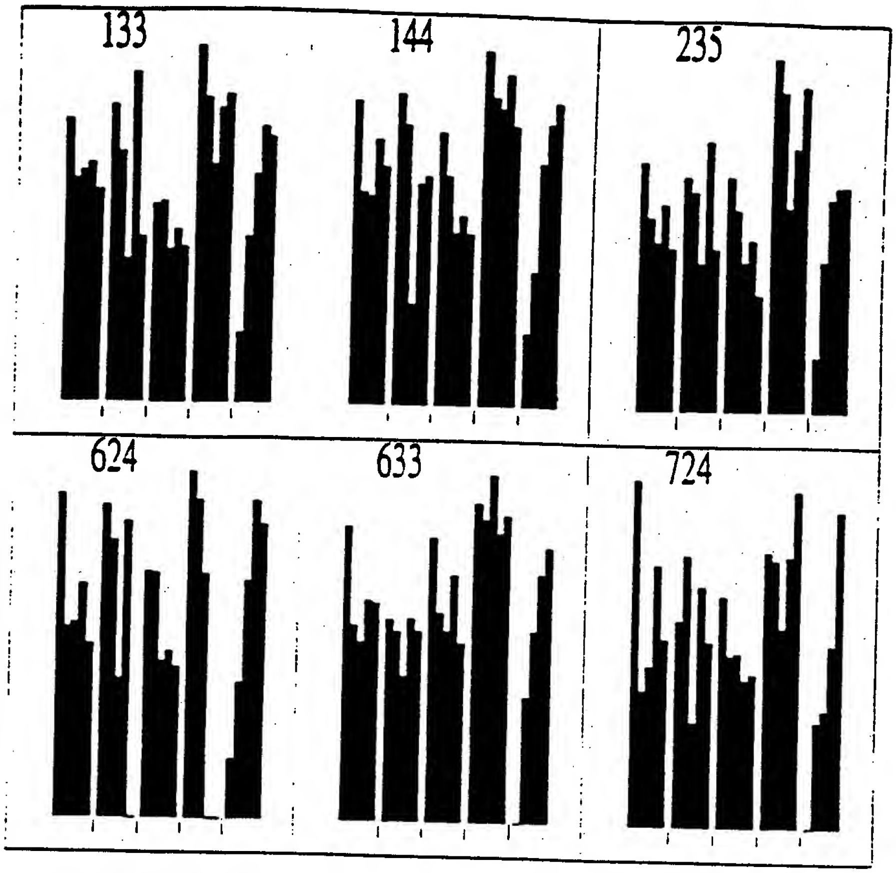


Figure 12. Data on a second coregulated group of spots, presented as in Fig. 11. The fourth experimental group (lovastatin) shows a modest induction, while the fifth group (lovastatin plus cholestyramine) does not.

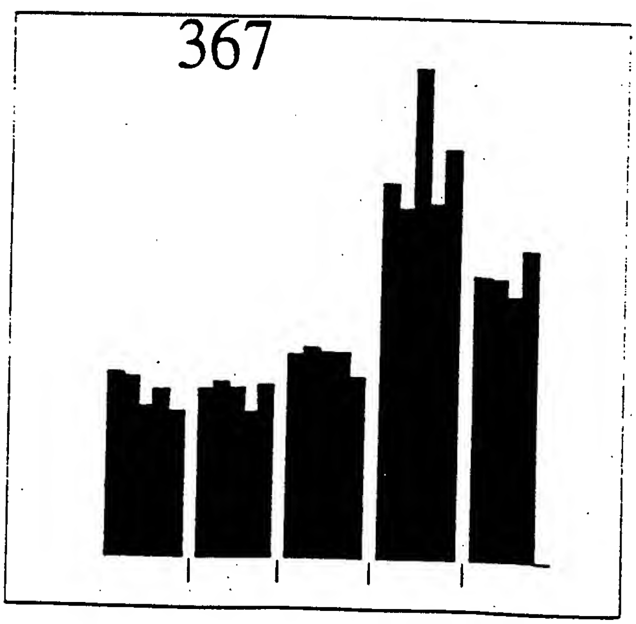


Figure 13. Data on spot MSN:367, presented as in Fig. 11. This protein shows unambiguously the anti-synergistic effect of lovastatin and cholestyramine (fifth group) as compared to lovastatin (fourth group). This response contrasts strongly with the regulation pattern seen in Fig. 11.

Addendum 2: Tables 1–4

Table 1. Master table of proteins in the rat liver database[a]

MSN	X	Y	CPKpI	SDSMW
3	311	434	<-35.0	63,800
5	568	263	-24.3	102,900
8	812	426	-16.0	64,800
11	549	268	-25.2	101,000
15	845	520	-15.3	55,200
17	629	589	-21.6	50,000
18	906	414	-14.0	66,300
19	755	298	-17.5	90,200
20	649	403	-20.9	67,900
21	1204	448	-8.7	62,100
22	332	434	<-35.0	63,800
23	787	424	-16.6	65,000
24	313	417	<-35.0	66,000
25	807	516	-16.1	55,500
27	1184	524	-9.0	54,900
28	1263	446	-8.0	62,400
29	743	605	-17.8	49,000
30	768	112	-17.2	348,600
32	1216	417	-8.6	66,000
33	1145	445	-9.5	62,500
34	1037	555	-11.3	52,400
35	863	412	-14.9	66,600
36	712	606	-18.7	48,900
38	763	694	-17.3	43,800
39	304	470	<-35.0	59,800
41	1165	569	-9.2	51,400
42	684	607	-19.6	48,800
43	1318	589	-7.3	50,000
44	1924	362	-0.1	74,600
46	1203	586	-8.7	50,200
47	1391	447	-6.3	62,300
48	309	454	<-35.0	61,500
49	605	587	-22.5	50,100
50	621	535	-21.8	53,900
51	1113	522	-10.0	55,000
52	1820	499	-0.9	57,000
53	725	177	-18.3	170,800
54	2001	500	>0.0	56,900
55	722	830	-18.4	37,300
56	678	533	-19.8	54,100
57	1682	302	-2.5	89,000
58	1091	580	-10.3	50,600
59	1171	585	-9.2	50,300
60	1400	624	-6.2	47,800
61	1853	508	-0.6	56,200
62	1888	567	-0.4	51,500
65	735	297	-18.1	90,500
66	1263	312	-8.0	85,900
67	1252	407	-8.1	67,300
68	779	692	-16.8	43,900
69	1064	296	-10.8	90,800
71	656	589	-20.6	50,000
72	638	545	-21.2	53,100
73	1582	583	-3.6	50,400
74	1570	556	-3.8	52,300
75	1264	621	-8.0	48,000
76	1338	564	-7.0	51,800
77	1833	363	-0.8	74,400
78	1767	565	-1.5	51,700
79	925	738	-13.6	41,600
80	534	698	-26.1	43,600
81	1811	363	-1.0	74,500
82	1412	681	-6.0	44,500
83	1471	347	-5.0	77,500
84	1662	563	-2.7	51,800
85	1596	479	-3.4	58,900
86	1817	301	-0.9	89,100
87	516	1371	-27.0	17,400
88	1589	698	-3.5	43,600
89	1706	719	-2.2	42,500
90	651	329	-20.8	81,700
91	1415	710	-6.0	43,000
92	1773	545	-1.4	53,200
93	1338	446	-7.0	62,300
94	1708	696	-2.2	43,700
95	1119	536	-9.9	53,800
96	1731	756	-2.0	40,700
97	1033	566	-11.4	51,600
98	1406	565	-6.1	51,700
99	578	1149	-23.8	25,000
100	2004	538	>0.0	53,700
101	1106	623	-10.1	47,900
102	482	455	-28.5	61,300
103	665	830	-20.2	37,300
104	773	1182	-17.0	23,800
105	312	1117	<-35.0	26,100
106	1769	509	-1.5	56,100
107	1585	720	-3.6	42,500
108	1692	807	-2.4	38,300
109	1482	593	-4.8	49,700
110	778	516	-16.9	55,500
111	1728	700	-2.0	43,500
113	1191	680	-8.9	44,500
114	1298	185	-7.5	160,800
115	682	907	-19.6	34,100
116	1146	610	-9.5	48,700
117	1548	849	-4.1	36,500
118	1050	577	-11.1	50,800
120	1530	828	-4.3	37,400
121	838	423	-15.4	65,200
122	1572	712	-3.8	42,900
123	23	1433	<-35.0	15,300
124	621	1474	-21.9	13,900
125	1298	862	-7.5	36,000
126	872	921	-14.7	33,500
127	1000	717	-12.0	42,600
128	1229	311	-8.4	86,100
129	1422	832	-5.8	37,300
130	1776	499	-1.4	57,000
131	1930	757	-0.1	40,700
132	660	537	-20.4	53,800
133	666	1019	-20.2	29,700
134	1271	862	-7.9	36,000
135	1161	1389	-9.3	16,800
136	453	1063	-29.7	28,100
137	1858	823	-0.6	37,700
138	1504	697	-4.6	43,700
139	1488	707	-4.8	43,200
140	1689	756	-2.4	40,700
141	311	1417	<-35.0	15,800
142	1366	915	-6.7	33,800
143	1429	346	-5.7	77,900
144	615	1017	-22.1	29,800
145	2006	566	>0.0	51,600
146	2006	518	>0.0	55,300
147	1070	1108	-10.7	26,500
148	1347	578	-6.9	50,800
149	541	1481	-25.7	13,700
150	1645	760	-2.8	40,500
151	1269	236	-7.9	117,000
152	1507	911	-4.5	33,900
153	1722	448	-2.1	62,100
154	932	503	-13.5	56,600
155	1031	294	-11.4	91,400
156	1970	684	>0.0	44,400
157	1258	183	-8.1	162,400
158	1275	417	-7.8	65,900
159	1663	820	-2.6	37,800
160	1034	527	-11.4	54,600
161	1953	771	>0.0	40,000
162	1020	1482	-11.6	13,700
164	1566	806	-3.8	38,400
166	1905	565	-0.2	51,700
167	1340	181	-7.0	164,900
168	1506	583	-4.6	50,400
169	1338	678	-7.0	44,700
170	1969	541	>0.0	53,500
171	800	378	-16.3	71,800
172	476	958	-28.7	32,100
173	919	1314	-13.7	19,300
174	1364	183	-6.7	162,900
175	825	393	-15.7	69,300
177	1582	553	-3.6	52,600
178	1321	710	-7.2	43,000
179	1089	615	-10.4	48,300
180	1866	567	-0.5	51,600
181	411	295	-32.1	91,200
182	804	730	-16.2	42,000
184	1860	896	-0.6	34,500
185	1997	1017	>0.0	29,800
186	279	1113	<-35.0	26,300
187	773	296	-17.0	90,800
188	1538	807	-4.2	38,400
191	1560	674	-3.9	44,900
192	1818	687	-0.9	44,200
193	1469	555	-5.0	52,400
194	1380	266	-6.4	101,600
195	784	632	-16.7	47,300
196	1227	1185	-8.4	23,700
197	667	553	-20.1	52,600
198	2006	681	>0.0	44,500
199	1711	674	-2.2	44,900
200	872	424	-14.7	65,000
201	292	435	<-35.0	63,700
202	736	253	-18.0	107,800
203	786	829	-16.7	37,400
204	1224	589	-8.5	50,000
205	439	983	-30.9	31,100
206	1994	571	>0.0	51,300
207	1895	687	-0.3	44,200
208	240	1418	<-35.0	15,800
210	1700	499	-2.3	57,000
211	902	517	-14.1	55,400
213	1087	684	-10.4	44,400
214	1340	668	-7.0	45,200
215	1591	495	-3.5	57,300
216	1585	755	-3.6	40,700
217	1159	393	-9.3	69,300
218	931	572	-13.5	51,200
219	713	177	-18.7	170,500
220	1479	911	-4.9	33,900
221	965	927	-12.8	33,300
223	934	716	-13.5	42,700
225	1812	1045	-1.0	28,800
226	821	411	-15.8	66,800
227	1586	1483	-3.6	13,600
228	1065	567	-10.8	51,600
229	1577	890	-3.7	34,800
230	1458	496	-5.2	57,300
232	1440	849	-5.5	36,500
234	1692	489	-2.4	57,900
235	618	1004	-22.0	30,300
236	920	1138	-13.7	25,400
237	952	1008	-13.1	30,200
238	1611	541	-3.2	53,500
239	1489	720	-4.8	42,500
240	501	448	-27.7	62,100
241	1820	569	-0.9	51,400
242	1357	658	-6.8	45,800
243	711	1182	-18.7	23,800
244	1855	621	-0.6	48,000
245	1189	474	-8.9	59,300
246	551	459	-25.1	61,000
247	1348	604	-6.9	49,100
248	460	448	-29.3	62,100
249	1733	451	-1.9	61,800
250	1974	788	>0.0	39,200
251	808	392	-16.1	69,500
252	874	553	-14.6	52,500
253	753	848	-17.6	36,500
254	995	450	-12.1	61,900
255	1690	679	-2.4	44,600
256	994	1006	-12.1	30,200
257	508	464	-27.4	60,400
258	1517	820	-4.4	37,800

[a] Master table of proteins in the rat liver database, showing spot master number, gel position (x and y), isoelectric point relative to CPK standards, and predicted molecular mass (from the standard curve of Fig. 8).

MSN	X	Y	CPKpI	SDSMW
259	1796	961	-1.1	31,900
260	661	1361	-20.4	17,700
261	1725	679	-2.0	44,600
262	496	1127	-28.0	25,800
263	1063	172	-10.9	177,400
265	1390	673	-6.3	45,000
266	510	437	-27.3	63,400
267	660	1038	-20.4	29,000
268	430	961	-31.0	31,900
269	1044	606	-11.2	48,900
270	2019	853	>0.0	36,300
271	857	422	-15.0	65,200
272	895	968	-14.2	31,700
274	1292	712	-7.6	42,900
275	1350	590	-6.9	49,900
276	1670	1089	-2.6	27,100
277	688	538	-19.4	53,700
278	961	718	-13.0	42,600
279	879	570	-14.5	51,300
281	1848	1084	-0.7	27,300
282	1505	525	-4.6	54,800
283	1313	1147	-7.3	25,100
284	1314	829	-7.3	37,400
285	1332	408	-7.1	67,200
286	1277	652	-7.8	46,100
288	1391	824	-6.3	37,600
289	1147	579	-9.5	50,700
290	925	511	-13.6	55,900
291	787	1476	-16.6	13,900
292	1462	818	-5.1	37,800
293	531	449	-26.3	62,000
294	860	698	-14.9	43,600
295	1162	609	-9.3	48,700
296	218	814	<-35.0	38,000
297	1377	979	-6.5	31,300
299	913	1523	-13.9	12,400
300	2012	667	>0.0	45,300
301	702	178	-19.0	169,200
302	494	1280	-28.1	20,400
303	403	1008	-32.6	30,100
304	1843	1585	-0.7	10,300
305	1049	593	-11.1	49,800
306	1608	989	-3.3	30,900
307	1219	916	-8.5	33,700
308	1627	755	-3.0	40,700
309	1524	892	-4.4	34,700
310	1769	1028	-1.5	29,400
311	1609	1451	-3.3	14,700
312	266	1408	<-35.0	16,100
313	1902	1365	-0.3	17,600
314	1316	1395	-7.3	16,600
315	1341	523	-7.0	54,900
318	1104	1053	-10.1	28,500
320	1480	1459	-4.9	14,400
321	850	603	-15.1	49,100
322	1454	1494	-5.3	13,300
323	670	626	-20.0	47,700
324	655	101	-20.6	420,500
325	1521	675	-4.4	44,800
326	1587	677	-3.6	44,700
327	1388	409	-6.3	67,000
328	448	1291	-30.0	20,100
330	1608	751	-3.3	40,900
331	1566	697	-3.8	43,700
332	531	471	-26.3	59,600
333	784	1156	-16.7	24,700
334	1059	407	-10.9	67,300
335	1593	303	-3.5	88,500
336	1616	598	-3.2	49,400
338	1854	1004	-0.6	30,300
339	1265	888	-8.0	34,900
340	581	585	-23.6	50,300
341	1497	1047	-4.7	28,700
343	1351	265	-6.8	102,200
344	1813	549	-0.9	52,800

MSN	X	Y	CPKpI	SDSMW
345	1006	578	-11.9	50,800
346	1095	640	-10.3	46,800
347	625	728	-21.7	42,000
348	361	983	-35.3	31,100
349	110	1343	<-35.0	18,300
350	521	1130	-26.7	25,700
351	912	619	-13.9	48,100
352	1574	530	-3.7	54,300
353	961	912	-12.9	33,900
354	706	762	-18.9	40,400
355	1450	830	-5.3	37,300
356	1374	1152	-6.5	24,900
357	474	997	-28.7	30,600
358	798	346	-16.3	77,800
359	764	338	-17.3	79,400
360	1384	1068	-6.4	27,900
361	1713	769	-2.1	40,100
362	1161	859	-9.3	36,100
363	914	1156	-13.8	24,800
364	412	435	-32.0	63,700
365	741	486	-17.9	58,200
366	878	1503	-14.6	13,000
367	1560	935	-3.9	33,000
368	983	520	-12.4	55,200
369	434	441	-31.0	63,000
370	639	610	-21.2	48,700
371	1587	860	-3.6	36,100
372	1875	762	-0.5	40,400
373	1351	1059	-6.8	28,300
374	1506	715	-4.6	42,700
375	1823	532	-0.9	54,200
376	254	417	<-35.0	65,900
377	1409	583	-6.1	50,400
378	621	494	-21.8	57,500
379	1017	595	-11.7	49,600
381	953	598	-13.1	49,400
382	856	674	-15.0	44,900
383	1252	258	-8.1	105,300
384	1699	1518	-2.3	12,500
385	1042	493	-11.2	57,500
386	1490	583	-4.7	50,400
387	1554	603	-4.0	49,100
388	1193	404	-8.9	67,700
389	1374	902	-6.5	34,300
390	1456	969	-5.2	31,700
391	718	690	-18.5	44,000
392	1799	732	-1.1	41,900
393	1482	758	-4.8	40,600
394	1227	1461	-8.4	14,400
395	1530	577	-4.3	50,800
396	1410	755	-6.0	40,800
397	912	256	-13.9	106,400
399	1465	1063	-5.0	28,100
400	1473	450	-4.9	61,900
401	1029	1140	-11.5	25,300
403	1516	754	-4.4	40,800
404	1495	554	-4.7	52,500
405	1525	1092	-4.3	27,100
406	723	252	-18.4	108,000
409	650	663	-20.8	45,500
410	1501	478	-4.6	59,000
411	936	1057	-13.4	28,300
412	350	1120	-35.9	26,000
413	1033	538	-11.4	53,700
415	737	425	-18.0	64,900
416	1578	606	-3.7	48,900
417	646	496	-21.0	57,300
418	1695	482	-2.3	58,600
419	725	770	-18.3	40,000
420	1289	1041	-7.7	28,900
421	1171	912	-9.1	33,900
422	599	162	-22.8	193,700
423	929	856	-13.6	36,200
424	739	625	-17.9	47,700
425	1490	965	-4.7	31,800

MSN	X	Y	CPKpI	SDSMW
426	1296	704	-7.6	43,300
427	810	843	-16.0	36,800
428	1565	303	-3.9	88,700
429	1259	847	-8.0	36,600
430	1253	562	-8.1	51,800
431	734	1426	-18.1	15,500
432	483	433	-28.5	63,900
434	518	1041	-26.9	28,900
435	1020	1170	-11.6	24,300
436	1122	196	-9.8	147,600
437	1870	673	-0.5	45,000
438	435	1102	-31.0	26,700
439	86	847	<-35.0	36,600
440	1740	544	-1.8	53,200
441	599	1571	-22.8	10,800
443	743	335	-17.8	80,100
446	801	668	-16.2	45,200
447	1050	926	-11.1	33,300
448	1245	1298	-8.2	19,800
449	1576	1516	-3.7	12,600
450	1818	1021	-0.9	29,600
451	1094	440	-10.3	63,100
452	1945	802	>0.0	38,600
453	1652	894	-2.8	34,600
454	1403	500	-6.1	56,900
456	1394	718	-6.3	42,600
457	905	436	-14.0	63,500
459	1038	581	-11.3	50,500
460	1598	294	-3.4	91,400
461	1528	863	-4.3	35,900
462	1098	1137	-10.2	25,400
463	849	1125	-15.2	25,800
464	1814	1072	-0.9	27,800
465	1388	481	-6.3	58,700
466	1194	1084	-8.9	27,300
468	577	467	-23.9	60,100
469	1140	888	-9.6	34,900
470	1797	524	-1.1	54,800
471	1293	1133	-7.6	25,500
472	618	655	-21.9	46,000
473	2009	299	>0.0	89,900
474	1205	215	-8.7	131,300
475	1035	788	-11.4	39,200
476	160	155	<-35.0	207,600
477	469	1370	-28.9	17,400
478	599	662	-22.8	45,600
479	1009	540	-11.8	53,500
480	1216	235	-8.6	117,400
482	816	346	-15.9	77,800
483	693	673	-19.3	44,900
485	1608	1013	-3.3	30,000
486	478	599	-28.6	49,300
487	1025	607	-11.5	48,800
488	1045	1186	-11.2	23,700
489	1609	301	-3.3	89,200
490	775	1289	-17.0	20,100
491	692	178	-19.3	169,300
492	1100	964	-10.2	31,800
493	1760	776	-1.6	39,700
494	882	247	-14.5	110,700
495	470	1258	-28.9	21,200
496	494	1436	-28.1	15,200
497	980	852	-12.5	36,400
499	1414	546	-6.0	53,100
500	1234	1072	-8.3	27,800
501	1246	659	-8.2	45,700
502	824	792	-15.7	39,000
503	1246	1134	-8.2	25,500
504	1115	1407	-9.9	16,200
505	1189	391	-8.9	69,700
506	1578	402	-3.7	68,000
507	787	250	-16.6	109,000
508	979	552	-12.5	52,800
509	1153	619	-9.4	48,100
510	1730	1006	-2.0	30,200

MSN	X	Y	CPKol	SDSMW
511	809	484	-16.0	58,400
512	1099	533	-10.2	54,100
513	1696	1034	-2.3	29,200
514	948	636	-13.2	47,100
515	481	543	-28.5	53,400
516	1334	1044	-7.1	28,800
517	868	1021	-14.8	29,700
518	798	779	-16.3	39,600
519	822	670	-15.7	45,100
520	632	165	-21.5	189,000
521	1332	830	-7.1	37,300
522	603	1104	-22.6	26,600
523	1190	309	-8.9	86,800
524	479	1226	-28.6	22,300
525	768	1066	-17.2	28,000
526	747	1016	-17.7	29,800
527	1170	231	-9.2	119,600
528	1502	542	-4.6	53,400
530	1728	620	-2.0	48,000
532	507	1011	-27.4	30,000
533	870	489	-14.7	57,900
534	1347	1085	-6.9	27,300
535	1513	346	-4.5	77,800
536	308	654	<-35.0	46,000
538	1851	689	-0.7	44,100
539	1463	982	-5.1	31,100
540	909	561	-13.9	52,000
541	625	289	-21.7	93,100
542	1164	198	-9.2	146,200
543	803	655	-16.2	45,900
544	1259	1143	-8.0	25,200
545	856	1526	-15.0	12,200
546	803	1071	-16.2	27,800
547	1162	274	-9.3	98,400
548	128	1321	<-35.0	19,000
549	1355	1122	-6.8	25,900
550	595	866	-23.0	35,800
552	1369	494	-6.6	57,500
553	992	405	-12.2	67,600
555	1125	410	-9.8	66,900
556	705	975	-18.9	31,400
557	1477	1030	-4.9	29,300
558	980	583	-12.5	50,400
559	700	1109	-19.1	26,400
560	1028	621	-11.5	48,000
562	898	794	-14.1	38,900
564	789	1446	-16.6	14,900
565	777	766	-16.9	40,200
566	980	328	-12.5	81,900
567	1519	611	-4.4	48,600
569	1212	661	-8.6	45,600
570	760	594	-17.4	49,700
571	618	956	-21.9	32,100
573	1142	771	-9.6	40,000
574	532	787	-26.2	39,300
575	771	250	-17.1	109,200
576	1068	534	-10.8	54,100
577	822	734	-15.7	41,800
578	914	754	-13.8	40,800
579	1064	794	-10.8	38,900
580	1524	714	-4.4	42,800
581	1392	783	-6.3	39,400
582	982	686	-12.4	44,200
584	1487	672	-4.8	45,000
585	758	731	-17.4	41,900
586	687	1152	-19.5	24,900
587	930	523	-13.5	55,000
588	1888	774	-0.4	39,900
589	642	485	-21.1	58,300
590	1317	519	-7.3	55,300
591	65	1548	<-35.0	11,500
592	1014	614	-11.7	48,400
593	732	176	-18.1	172,300
594	1627	478	-3.0	59,000
595	1009	1426	-11.8	15,500
596	619	269	-21.9	100,500
597	1176	461	-9.1	60,700
598	1465	1044	-5.0	28,800
599	741	1188	-17.9	23,600
600	907	402	-14.0	68,000
601	687	658	-19.5	45,800
602	712	1138	-18.7	25,400
603	898	181	-14.1	165,200
604	783	1461	-16.7	14,400
605	736	223	-18.0	125,300
606	629	273	-21.6	98,700
607	1064	286	-10.8	94,000
608	883	503	-14.5	56,700
609	2012	610	>0.0	48,700
610	1255	903	-8.1	34,200
612	1103	391	-10.1	69,600
613	778	265	-16.9	102,000
614	824	518	-15.7	55,400
615	1095	195	-10.3	149,100
616	1758	478	-1.6	59,000
617	994	372	-12.1	72,900
618	751	374	-17.6	72,400
619	1429	518	-5.7	55,300
620	1050	520	-11.1	55,200
621	923	1105	-13.7	26,600
622	1462	622	-5.1	47,900
623	759	225	-17.4	124,000
624	758	1038	-17.4	29,000
625	1438	606	-5.5	48,900
626	1096	1089	-10.2	27,200
627	942	548	-13.3	53,000
628	809	621	-16.0	48,000
629	899	979	-14.1	31,300
630	1135	1321	-9.6	19,100
631	979	615	-12.5	48,300
632	1542	1076	-4.1	27,600
633	1345	814	-6.9	38,000
634	409	950	-32.2	32,400
635	1165	704	-9.2	43,300
636	774	604	-17.0	49,000
637	1263	524	-8.0	54,800
638	952	411	-13.1	66,700
639	1717	575	-2.1	51,000
640	994	292	-12.1	92,000
641	165	1224	<-35.0	22,400
642	803	251	-16.2	108,900
643	719	296	-18.5	90,700
644	1100	294	-10.2	91,400
645	534	1263	-26.1	21,000
646	1153	1038	-9.4	29,000
648	1246	204	-8.2	140,000
649	14	1406	<-35.0	16,200
650	1713	1049	-2.1	28,600
651	1986	1183	>0.0	23,800
652	1378	816	-6.5	38,000
653	1442	1165	-5.5	24,400
654	650	806	-20.8	38,400
655	1111	551	-10.0	52,700
656	1095	861	-10.3	36,000
657	1524	540	-4.4	53,600
658	1777	860	-1.4	36,000
659	391	584	-33.4	50,400
660	977	565	-12.5	51,700
661	658	166	-20.5	187,500
662	732	312	-18.1	86,100
663	1787	567	-1.2	51,500
664	888	268	-14.4	100,900
665	889	775	-14.3	39,800
666	715	221	-18.6	126,300
667	781	227	-16.8	122,400
668	646	165	-21.0	189,100
669	1116	353	-9.9	76,300
670	1382	643	-6.4	46,600
671	547	789	-25.3	39,200
673	984	746	-12.4	41,200
674	1661	448	-2.7	62,100
675	1523	562	-4.4	51,900
676	708	642	-18.8	46,700
677	919	615	-13.7	48,300
678	1085	551	-10.5	52,700
679	600	923	-22.7	33,400
680	1237	1004	-8.3	30,300
681	1103	283	-10.1	95,100
682	1406	477	-6.1	59,100
683	1596	249	-3.4	109,800
684	555	699	-24.8	43,500
685	1167	1313	-9.2	19,300
686	1932	790	0.0	39,100
687	1545	619	-4.1	48,100
688	1456	764	-5.2	40,300
689	1011	953	-11.8	32,300
690	1995	270	>0.0	100,200
691	812	888	-16.0	34,900
692	1154	1461	-9.4	14,400
693	1993	819	>0.0	37,800
694	1628	656	-3.0	45,900
695	928	254	-13.6	107,000
696	1854	715	-0.6	42,700
697	1997	345	>0.0	78,000
698	957	563	-13.0	51,800
699	1540	730	-4.2	42,000
702	577	900	-23.8	34,400
703	1610	562	-3.2	51,900
705	1278	571	-7.8	51,200
706	1841	704	-0.7	43,300
707	1018	1386	-11.7	16,900
709	1074	1145	-10.7	25,100
710	293	889	<-35.0	34,800
712	720	412	-18.5	66,600
713	1386	841	-6.4	36,800
714	1328	263	-7.1	103,100
715	698	433	-19.1	63,900
716	701	481	-19.0	58,700
717	1875	699	-0.5	43,600
718	575	702	-23.9	43,400
719	1216	204	-8.6	140,400
721	1069	464	-10.8	60,400
722	1272	506	-7.9	56,400
723	958	822	-13.0	37,700
724	763	395	-17.3	69,100
725	720	916	-18.5	33,700
726	1476	415	-4.9	66,200
727	1846	473	-0.7	59,400
728	510	783	-27.3	39,400
729	1217	1126	-8.6	25,800
730	1858	724	-0.6	42,300
731	665	765	-20.2	40,300
733	1321	312	-7.2	85,900
734	719	427	-18.5	64,600
735	1101	473	-10.2	59,500
736	1359	569	-6.7	51,400
738	696	220	-19.2	127,600
739	687	409	-19.5	67,000
740	1205	256	-8.7	106,200
741	995	563	-12.1	51,900
742	898	596	-14.1	49,500
743	881	181	-14.5	165,900
744	1951	686	>0.0	44,200
745	726	168	-18.3	183,600
746	999	643	-12.0	46,600
748	182	1503	<-35.0	13,000
749	2005	649	>0.0	46,300
750	1448	575	-5.4	51,000
751	792	266	-16.5	101,900
752	469	296	-28.9	90,600
754	664	254	-20.3	107,000
755	1195	184	-8.8	161,000
756	1821	1113	-0.9	26,300
757	909	246	-13.9	111,000
760	790	133	-16.5	264,900

MSN	X	Y	CPKpI	SDSMW
761	1399	733	-6.2	41,800
763	1416	1085	-5.9	27,300
764	2020	569	>0.0	51,400
765	651	475	-20.8	59,300
766	1052	1149	-11.1	25,000
767	1968	468	>0.0	59,900
768	1330	685	-7.1	44,300
769	1970	613	>0.0	48,500
770	857	617	-15.0	48,200
771	1337	974	-7.0	31,500
773	1576	502	-3.7	56,700
775	969	824	-12.8	37,600
776	1438	708	-5.5	43,100
777	1539	458	-4.2	61,000
778	850	434	-15.1	63,800
779	700	411	-19.1	66,800
780	1052	1136	-11.1	25,500
784	1413	529	-6.0	54,400
785	1364	885	-6.7	35,000
786	1822	835	-0.9	37,100
787	893	392	-14.3	69,500
790	616	882	-22.0	35,100
791	451	1429	-29.8	15,400
792	777	377	-16.9	72,000
793	1536	1543	-4.2	11,700
794	1461	807	-5.1	38,300
796	388	546	-33.6	53,100
797	1126	212	-9.8	133,700
798	933	437	-13.5	63,400
799	1420	593	-5.9	49,800
800	1759	279	-1.6	96,500
801	624	865	-21.7	35,800
802	898	547	-14.2	53,000
803	1775	1468	-1.4	14,200
804	573	196	-24.0	148,400
805	203	494	<-35.0	57,400
806	980	1039	-12.5	29,000
807	902	308	-14.1	87,200
808	625	827	-21.7	37,500
809	1851	1015	-0.7	29,900
810	440	573	-30.9	51,100
811	1358	249	-6.8	109,700
812	851	393	-15.1	69,400
813	745	1246	-17.8	21,600
814	2028	810	>0.0	38,200
815	1086	645	-10.4	46,500
816	629	313	-21.6	85,700
817	1376	1177	-6.5	24,000
818	1771	790	-1.4	39,100
819	1045	263	-11.2	103,100
820	984	362	-12.4	74,600
821	1712	279	-2.2	96,700
822	1256	205	-8.1	139,200
823	1517	654	-4.4	46,000
824	1442	449	-5.5	62,000
825	1240	513	-8.3	55,800
826	1309	1014	-7.4	29,900
827	2012	708	>0.0	43,100
828	937	1405	-13.4	16,200
830	1342	756	-7.0	40,700
831	562	826	-24.5	37,500
832	1073	1039	-10.7	29,000
833	481	820	-28.5	37,800
834	501	581	-27.8	50,500
837	751	748	-17.6	41,100
838	635	833	-21.3	37,200
839	1494	459	-4.7	60,900
840	1952	301	>0.0	89,300
841	1585	1080	-3.6	27,500
842	571	1312	-24.1	19,400
843	1325	649	-7.2	46,300
844	1727	301	-2.0	89,200
845	630	679	-21.5	44,600
846	2016	905	>0.0	34,200
847	673	1200	-19.9	23,200
848	1863	271	-0.6	99,500
849	1166	523	-9.2	54,900
850	1535	1024	-4.2	29,600
851	1035	826	-11.4	37,500
852	834	542	-15.5	53,400
855	499	220	-27.8	127,100
856	1063	194	-10.9	150,500
857	887	890	-14.4	34,800
858	1448	639	-5.4	46,900
859	706	311	-18.9	86,200
860	1070	1066	-10.7	28,000
861	472	347	-28.8	77,600
862	674	480	-19.9	58,800
864	1307	499	-7.4	57,000
865	645	887	-21.0	34,900
866	827	1004	-15.6	30,300
868	685	494	-19.5	57,400
869	1807	402	-1.0	68,000
870	1323	783	-7.2	39,400
871	1228	1031	-8.4	29,300
872	1904	346	-0.3	77,700
873	556	647	-24.8	46,400
874	1540	756	-4.2	40,700
875	1566	777	-3.8	39,700
876	1198	351	-8.8	76,800
877	1076	720	-10.6	42,500
878	1161	1111	-9.3	26,400
879	647	757	-20.9	40,700
880	1756	594	-1.6	49,700
881	1543	278	-4.1	97,100
883	1432	890	-5.7	34,800
884	922	689	-13.7	44,100
885	1103	414	-10.1	66,400
886	1501	607	-4.6	48,900
887	798	1103	-16.3	26,600
888	636	634	-21.3	47,200
889	951	759	-13.1	40,600
890	717	548	-18.6	52,900
891	1123	229	-9.8	121,200
892	891	413	-14.3	66,400
894	1245	234	-8.2	117,800
895	1962	346	>0.0	77,700
896	1322	626	-7.2	47,700
897	420	570	-31.4	51,300
898	662	428	-20.3	64,500
899	845	243	-15.3	113,000
900	624	703	-21.7	43,400
901	931	1094	-13.5	27,000
903	799	229	-16.3	121,000
904	765	520	-17.2	55,200
905	775	889	-17.0	34,800
907	888	824	-14.4	37,600
908	828	1303	-15.6	19,700
910	681	1544	-19.7	11,700
911	1544	301	-4.1	89,100
913	1606	387	-3.3	70,400
914	1237	688	-8.3	44,100
916	1442	749	-5.5	41,100
917	1260	367	-8.0	73,700
919	764	1541	-17.3	11,700
920	1133	1123	-9.7	25,900
921	1123	380	-9.8	71,500
923	829	242	-15.6	113,200
924	1131	318	-9.7	84,300
925	1441	874	-5.5	35,400
926	679	219	-19.7	128,200
927	1487	1191	-4.8	23,500
928	1082	775	-10.5	39,800
929	1231	816	-8.4	38,000
931	1609	670	-3.3	45,100
932	810	900	-16.0	34,400
933	965	520	-12.8	55,100
934	947	462	-13.2	60,600
936	865	843	-14.8	36,800
937	1421	1056	-5.9	28,400
939	1197	827	-8.8	37,500
941	1765	885	-1.5	35,000
942	602	472	-22.7	59,600
943	312	498	<-35.0	57,100
944	993	491	-12.1	57,700
945	1300	269	-7.5	100,300
946	630	423	-21.6	65,100
947	187	736	<-35.0	41,600
948	1380	344	-6.5	78,200
949	1766	665	-1.5	45,400
950	1038	193	-11.3	151,000
951	860	152	-14.9	213,000
952	957	701	-13.0	43,400
954	503	547	-27.6	53,000
955	1938	712	>0.0	42,900
957	1010	816	-11.8	37,900
959	768	174	-17.2	174,900
960	596	419	-23.0	65,700
961	557	409	-24.8	67,100
962	887	320	-14.4	83,900
963	564	334	-24.5	80,500
964	969	1155	-12.8	24,800
965	671	255	-20.0	106,600
966	1204	798	-8.7	38,700
967	910	154	-13.9	210,300
968	609	1048	-22.3	28,700
969	1285	206	-7.7	138,900
970	822	232	-15.8	119,300
971	976	437	-12.6	63,400
972	403	567	-32.6	51,600
974	279	495	<-35.0	57,400
975	844	981	-15.3	31,200
976	1124	295	-9.8	91,100
977	994	664	-12.1	45,400
978	1612	642	-3.2	46,700
979	749	1141	-17.7	25,300
980	1064	642	-10.8	46,700
981	1197	911	-8.8	33,900
983	1762	1508	-1.6	12,800
984	1344	317	-6.9	84,700
985	1024	1105	-11.5	26,600
987	739	1159	-17.9	24,600
988	816	555	-15.9	52,400
990	785	361	-16.7	74,900
991	1159	317	-9.3	84,500
992	1090	928	-10.4	33,300
993	1030	701	-11.5	43,400
994	847	811	-15.2	38,200
995	902	461	-14.1	60,700
996	888	847	-14.4	36,600
997	1815	579	-0.9	50,700
998	1205	504	-8.7	56,500
999	617	289	-22.0	93,100
1000	968	290	-12.8	92,700
1001	970	771	-12.7	40,000
1002	1736	478	-1.9	58,900
1003	643	1184	-21.1	23,700
1006	822	487	-15.8	58,100
1007	875	279	-14.6	96,400
1009	291	644	<-35.0	46,600
1010	1386	745	-6.4	41,200
1011	459	541	-29.4	53,500
1012	679	661	-19.7	45,600
1013	1818	1128	-0.9	25,800
1014	1032	634	-11.4	47,200
1015	1629	994	-3.0	30,700
1016	1311	1134	-7.4	25,500
1017	1722	424	-2.0	65,000
1018	1015	743	-11.7	41,300
1020	1574	1219	-3.7	22,500
1021	781	484	-16.8	58,400
1022	1129	83	-9.7	591,300
1023	812	317	-15.9	84,600
1024	785	446	-16.7	62,400
1025	1290	739	-7.7	41,500

MSN	X
[illegible]	405
[illegible]	1296
[illegible]	856
[illegible]	1284
[illegible]	986
[illegible]	1547
[illegible]	1381
[illegible]	1525
[illegible]	1128
[illegible]	1226
[illegible]	1761
[illegible]	541
[illegible]	818
[illegible]	1036
[illegible]	1439
[illegible]	1540
[illegible]	1576
[illegible]	1089
[illegible]	949
[illegible]	426
[illegible]	1583
[illegible]	779
[illegible]	1613
[illegible]	1380
[illegible]	284
[illegible]	1261
[illegible]	393
[illegible]	1817
[illegible]	1245
[illegible]	1258
[illegible]	705
[illegible]	1181
[illegible]	529
[illegible]	508
[illegible]	1898
[illegible]	873
[illegible]	1768
[illegible]	836
[illegible]	1863
[illegible]	826
[illegible]	971
[illegible]	1697
[illegible]	1157
[illegible]	620
[illegible]	1867
[illegible]	2019
[illegible]	1546
[illegible]	1545
[illegible]	61
[illegible]	1954
[illegible]	588
[illegible]	1050
[illegible]	457
[illegible]	1884
[illegible]	1714
[illegible]	1717
[illegible]	1976
[illegible]	547
[illegible]	1348
[illegible]	1385
[illegible]	1078
[illegible]	975
[illegible]	1202
[illegible]	1022
[illegible]	1905
[illegible]	1512
[illegible]	1114
[illegible]	1464
[illegible]	1048
[illegible]	1122
[illegible]	1722
[illegible]	1098
[illegible]	1830
[illegible]	764
[illegible]	1968

MSN	X	Y	CPKpI	SDSMW
1026	405	552	-32.3	52,600
1027	1298	848	-7.5	36,500
1028	856	547	-15.0	53,000
1030	1284	226	-7.7	123,200
1031	986	822	-12.3	37,700
1032	1547	403	-4.1	67,900
1033	1381	551	-6.4	52,700
1034	1525	496	-4.3	57,200
1035	1128	645	-9.7	46,500
1036	1226	274	-8.5	98,300
1039	1761	262	-1.6	103,600
1040	541	839	-25.7	36,900
1041	818	910	-15.8	34,000
1044	1036	485	-11.3	58,300
1045	1439	407	-5.5	67,300
1047	1540	250	-4.2	109,200
1048	1576	635	-3.7	47,100
1049	1089	411	-10.4	66,700
1050	949	1040	-13.2	28,900
1051	426	818	-31.1	37,800
1052	1583	1385	-3.6	16,900
1053	779	1092	-16.8	27,000
1054	1613	620	-3.2	48,000
1055	1380	377	-6.5	72,000
1056	284	663	<-35.0	45,500
1058	1261	746	-8.0	41,200
1060	393	605	-33.3	49,000
1061	1817	645	-0.9	46,600
1062	1245	746	-8.2	41,200
1064	1258	792	-8.1	39,000
1065	705	934	-18.9	33,000
1066	1181	734	-9.0	41,800
1067	529	658	-26.3	45,800
1068	508	696	-27.4	43,700
1069	1898	604	-0.3	49,100
1071	873	609	-14.7	48,700
1073	1768	1128	-1.5	25,800
1075	836	773	-15.4	39,900
1076	1863	861	-0.6	36,000
1078	826	566	-15.7	51,600
1081	971	483	-12.7	58,500
1083	1697	202	-2.3	142,300
1085	1157	794	-9.4	38,900
1090	620	910	-21.9	34,000
1092	1867	597	-0.5	49,500
1093	2019	894	>0.0	34,600
1094	1546	538	-4.1	53,700
1095	1545	477	-4.1	59,100
1098	61	935	<-35.0	33,000
1099	1954	237	>0.0	116,000
1101	588	1048	-23.3	28,600
1102	1050	667	-11.1	45,200
1103	457	797	-29.5	38,800
1105	1884	532	-0.4	54,200
1106	1714	649	-2.1	46,300
1107	1717	546	-2.1	53,100
1108	1976	722	>0.0	42,400
1111	547	1066	-25.3	28,000
1112	1348	621	-6.9	48,000
1115	1385	762	-6.4	40,400
1116	1078	816	-10.6	38,000
1117	975	787	-12.6	39,300
1118	1202	933	-8.7	33,100
1119	1022	1076	-11.6	27,600
1120	1905	616	-0.3	48,300
1121	1512	1301	-4.5	19,700
1122	1114	677	-9.9	44,700
1123	1464	452	-5.1	61,700
1125	1048	857	-11.1	36,200
1126	1122	802	-9.8	38,600
1128	1722	892	-2.1	34,700
1133	1098	825	-10.2	37,500
1139	1830	569	-0.8	51,400
1147	764	1182	-17.3	23,800
1148	1968	724	>0.0	42,300

MSN	X	Y	CPKpI	SDSMW
1153	921	1158	-13.7	24,700
1154	1594	864	-3.5	35,900
1161	637	400	-21.3	68,400
1162	623	397	-21.8	68,800
1163	665	397	-20.2	68,700
1168	564	528	-24.4	54,500
1170	552	529	-25.0	54,500
1171	538	524	-25.9	54,800
1172	545	514	-25.5	55,700
1174	1099	522	-10.2	55,000
1176	1304	586	-7.5	50,200
1177	1366	539	-6.6	53,700
1178	1608	702	-3.3	43,400
1179	1485	224	-4.8	124,900
1180	1459	224	-5.2	124,900
1181	1431	223	-5.7	125,100
1182	1407	223	-6.1	125,200
1183	1383	224	-6.4	124,700
1184	1454	182	-5.3	164,400
1185	1422	183	-5.8	162,600
1186	1394	182	-6.3	164,300
1189	1171	214	-9.2	131,800
1190	1457	286	-5.2	94,200
1191	686	1114	-19.5	26,200
1192	265	893	<-35.0	34,700
1193	403	1292	-32.6	20,000
1194	344	1275	<-35.0	20,600
1195	505	1311	-27.6	19,400
1196	572	1293	-24.1	20,000
1197	639	1502	-21.2	13,000
1198	637	1402	-21.3	16,300
1199	614	1407	-22.1	16,200
1200	637	1431	-21.3	15,400
1201	1095	1394	-10.3	16,600
1202	1719	1545	-2.1	11,600
1203	791	668	-16.5	45,200
1204	964	1021	-12.9	29,700
1205	313	195	<-35.0	148,700
1208	306	194	<-35.0	149,800
1209	320	197	<-35.0	147,400
1210	326	197	<-35.0	146,600
1211	394	294	-33.2	91,400
1212	402	294	-32.7	91,200
1214	386	294	-33.7	91,400
1215	641	329	-21.2	81,600
1216	660	329	-20.4	81,600
1217	914	266	-13.8	101,800
1218	873	245	-14.7	112,000
1219	970	372	-12.7	72,900
1220	1021	298	-11.6	90,100
1221	1392	205	-6.3	139,500
1222	1354	203	-6.8	141,800
1223	1362	205	-6.7	139,500
1224	673	540	-19.9	53,600
1225	614	542	-22.1	53,400
1226	603	539	-22.6	53,600
1227	686	623	-19.2	47,800
1228	707	628	-18.9	47,500
1229	475	447	-28.7	62,300
1230	466	1282	-29.0	20,400
1231	759	1461	-17.4	14,400
1232	1324	1170	-7.2	24,200
1233	1583	1005	-3.6	30,300
1234	1865	809	-0.6	38,200
1235	1812	817	-1.0	37,900
1236	1411	703	-6.0	43,400
1237	1392	682	-6.3	44,500
1238	794	410	-16.4	66,900
1239	769	407	-17.1	67,300
1240	740	406	-17.9	67,500
1241	743	511	-17.8	55,900
1242	713	510	-18.7	56,000
1243	682	509	-19.6	56,100
1244	663	504	-20.3	56,500
1245	565	582	-24.4	50,500

MSN	X	Y	CPKpI	SDSMW
1246	547	577	-25.3	50,800
1247	530	576	-26.3	50,900
1249	516	572	-27.0	51,200
1250	973	536	-12.7	53,900
1251	607	532	-22.4	54,200
1252	665	529	-20.2	54,400
1253	899	766	-14.1	40,200
1254	1311	746	-7.4	41,200
1255	1300	761	-7.5	40,400
1257	1938	712	0.0	42,900
1258	1806	718	-1.0	42,600
1259	1727	715	-2.0	42,700
1260	1629	713	-3.0	42,800
1261	1555	717	-4.0	42,600
1262	1468	717	-5.0	42,600
1263	1413	722	-6.0	42,400
1264	1340	717	-7.0	42,600
1265	1263	717	-8.0	42,600
1266	1182	720	-9.0	42,500
1267	1110	717	-10.0	42,600
1268	1055	717	-11.0	42,600
1269	999	717	-12.0	42,600
1270	959	715	-13.0	42,700
1271	905	712	-14.0	42,900
1272	857	714	-15.0	42,800
1273	810	705	-16.0	43,300
1274	774	711	-17.0	42,900
1277	737	708	-18.0	43,100
1278	702	711	-19.0	42,900
1279	671	710	-20.0	43,000
1280	645	710	-21.0	43,000
1281	617	707	-22.0	43,100
1282	595	704	-23.0	43,300
1283	573	700	-24.0	43,500
1284	552	695	-25.0	43,700
1285	536	694	-26.0	43,800
1286	515	687	-27.0	44,200
1287	496	683	-28.0	44,400
1288	467	669	-29.0	45,200
1289	447	667	-30.9	45,300
1290	427	655	-31.0	45,900
1291	412	655	-32.0	45,900
1292	397	652	-33.0	46,100
1293	381	654	-34.0	46,000
1294	365	653	-35.0	46,100
1295	348	653	<-35.0	46,100

Table 2. Table of some identified proteins

POP name	Protein name	MSN's	Basis for identification
IDS:3_ALPHA_HDDH	3-α-hydroxysteroid-dihydrodiol-dehydrogenase, an enzyme of steroid metabolism	137, 159	Pure protein and antibody provided by Dr. T.M. Penning, Department of Pharmacology, School of Medicine, University of Pennsylvania.
IDS:ACTIN_BETA	β cellular actin, a cytoskeletal protein	38	Homologous position with respect to other mammalian systems
IDS:ACTIN_GAMMA	γ cellular actin, a cytoskeletal protein	68	Homologous position with respect to other mammalian systems
IDS:ALBUMIN	Serum albumin, mature form.	21, 28, 33	Predominance in rat plasma
IDS:APO_A-I	Apo A-I plasma lipoprotein, mature form (tentative).	236, 463	Presence in rat plasma, regulation by some lipid-lowering drugs
IDS:CALMODULIN	Calmodulin, an acidic cytosolic calcium-binding protein	123, 649	Homologous position with respect to other mammalian systems
IDS:CATALASE	Catalase (peroxisomal)	54, 61, 106	Presence in purified peroxisomes, similarity in position to mouse catalase
IDS:CPKSPOTS	Spots contributed by the CPK charge standards (not rat liver proteins)	1257 - 1295	
IDS:CPS	Carbamoyl phosphate synthase	114, 157, 167, 174, 1184, 1185, 1186, 1222	Pure protein provided by Dr. Margaret Marshall, Department of Pharmacology, Medical School, University of Wisconsin - Madison.
IDS:CYTOCHROME_B5	Cytochrome b5	87, 477	Pure protein provided by Dr. Andrew Parkinson, Department of Pharmacology, Toxicology and Therapeutics, University of Kansas Medical Center
IDS:FABP-L	Liver fatty-acid binding protein	227	Pure protein provided by Dr. Nathan Bass, Department of Medicine, University of California School of Medicine, San Francisco
IDS:HMG-COA_SYNTHASE	Cytosolic HMG-CoA Synthase	133, 144, 235, 413	Antibody provided by Dr. Michael Greenspan, Merck Sharp & Dohme Research Laboratories, Rahway, NJ
IDS:LAMIN_B	Lamin B, a nuclear protein	415, 734	Homologous position with respect to other mammalian systems
IDS:MITCON:1	Mitcon:1 (F1 ATPase β subunit), a mitochondrial inner membrane	17, 49, 71, 340, 1245, 1246, 1247, 1249	Homologous position with respect to other mammalian systems, presence in mitochondria
IDS:MITCON:2	Mitcon:2, a mitochondrial matrix stress protein equivalent to E.	15, 25, 110, 1241, 1242, 1243, 1244	Homologous position with respect to other mammalian systems, presence in mitochondria
IDS:MITCON:3	Mitcon:3, a mitochondrial matrix stress protein, likely analog of	18, 35, 226, 600, 1238, 1239, 1240	Homologous position with respect to other mammalian systems, presence in mitochondria
IDS:NADPH_P450_RED	NADPH cytochrome P-450 reductase, frequently co-induced with P-450's	175, 251, 812	Pure protein provided by Dr. Andrew Parkinson, Department of Pharmacology, Toxicology and Therapeutics, University of Kansas Medical Center
IDS:PDI	Protein disulphide isomerase 1	168, 1170, 1171, 1172	Sequence information obtained by R.M. Van Frank, Lilly Research Laboratories, Indianapolis
IDS:PLASMA_PROTEINS	Rat plasma proteins observed in liver	21, 28, 33, 44, 72, 102, 115, 197, 236, 246, 248, 257, 293, 332, 347, 364, 369, 419, 432, 463, 468, 518, 562, 605, 623, 666, 667, 725, 738, 790, 865, 903, 926	Plasma coelectrophoresis studies
IDS:PRO-ALBUMIN	Serum albumin precursor	47, 93	Relative position to mature albumin, presence in microsomes
IDS:PYRCARBOX	Pyruvate carboxylase	179, 1180, 1181, 1182, 1183	Pavlica, R.J., et al., BBA (1990) *1022* 115-125.
IDS:SOD	Superoxide dismutase	135	Sequence information obtained by R.M. Van Frank, Lilly Research Laboratories, Indianapolis
IDS:TUBULIN_ALPHA	α tubulin, a cytoskeletal protein	56, 132, 1224, 1252	Homologous position with respect to other mammalian systems
IDS:TUBULIN_BETA	β tubulin, a cytoskeletal protein	50, 1225, 1226, 1251	Homologous position with respect to other mammalian systems

…e 3. **Computed pI's of two sets of carbamylated protein standards: Rabbit muscle CPK and human hemoglobin (Hb)**

Protein Name	PIR Name	#ASP 3.9	#GLU 4.1	#HIS 6.0	#LYS 10.8	#ARG 12.5	NH2- 7.0	Calc pI	Real CPK
Rabbit muscle CPK	KIRBCM	28	27	17	34	18	1	6.84	0.0
		28	27	17	33	18	1	6.67	-1
		28	27	17	32	18	1	6.54	-2
		28	27	17	31	18	1	6.42	-3
		28	27	17	30	18	1	6.31	-4
		28	27	17	29	18	1	6.21	-5
		28	27	17	28	18	1	6.12	-6
		28	27	17	27	18	1	6.03	-7
		28	27	17	26	18	1	5.94	-8
		28	27	17	25	18	1	5.85	-9
		26	27	17	24	18	1	5.76	-10
		28	27	17	23	18	1	5.67	-11
		28	27	17	22	18	1	5.58	-12
		28	27	17	21	18	1	5.48	-13
		28	27	17	20	18	1	5.39	-14
		28	27	17	19	18	1	5.29	-15
		28	27	17	18	18	1	5.20	-16
		28	27	17	17	18	1	5.12	-17
		28	27	17	16	18	1	5.04	-18
		28	27	17	15	18	1	4.96	-19
		28	27	17	14	18	1	4.89	-20
		28	27	17	13	18	1	4.83	-21
		28	27	17	12	18	1	4.77	-22
		28	27	17	11	18	1	4.71	-23
		28	27	17	10	18	1	4.66	-24
		28	27	17	9	18	1	4.61	-25
		28	27	17	8	18	1	4.56	-26
		28	27	17	7	18	1	4.52	-27
		28	27	17	6	18	1	4.48	-28
		28	27	17	5	18	1	4.44	-29
		28	27	17	4	18	1	4.40	-30
		28	27	17	3	18	1	4.36	-31
		28	27	17	2	18	1	4.32	-32
		28	27	17	1	18	1	4.29	-33
		28	27	17	0	18	1	4.25	-34
		28	27	17	0	18	0	4.22	-35
Hb-beta, human	HBHU	7	8	9	11	3	1	7.18	
		7	8	9	10	3	1	6.79	
		7	8	9	9	3	1	6.53	-1.8
		7	8	9	8	3	1	6.32	-3.2
		7	8	9	7	3	1	6.13	-5.3
		7	8	9	6	3	1	5.96	-7.2
		7	8	9	5	3	1	5.78	-10.0
		7	8	9	4	3	1	5.59	-12.3
		7	8	9	3	3	1	5.37	-15.5
		7	8	9	2	3	1	5.14	-18.0
		7	8	9	1	3	1	4.91	-21.0
		7	8	9	0	3	1	4.71	-25.5
		7	8	9	0	3	0	4.54	-27.2

Table 4. Computed p*I*'s of some known proteins related to measured CPK p*I*'s

Protein Name	PIR Name	#ASP 3.9	#GLU 4.1	#HIS 6.0	#LYS 10.8	#ARG 12.5	Calc pI	Real CPK
Creatine phospho kinase (CPK), rabbit muscle	KIRBCM	28	27	17	34	18	6.84	0.0
Fatty acid-binding protein, rat hepatic	FZRTL	5	13	2	16	2	7.83	-3.0
b2-microglobulin, human	MGHUB2	7	8	4	8	5	6.09	-5.0
Carbamoyl-phosphate synthase, rat	SYRTCA	72	96	28	95	56	5.97	-5.5
Proalbumin (serum albumin precursor), rat	ABRTS	32	57	15	53	27	5.98	-6.2
Serum albumin, rat	ABRTS	32	57	15	53	24	5.71	-9.0
Superoxid dismutase (Cu-Zn, SOD), rat	A26810	8	11	10	9	4	5.91	-9.2
Phospholipase C, phophoinositide-specific (?), rat	A28807	34	42	9	49	21	5.92	-9.2
Albumin, human	ABHUS	36	61	16	60	24	5.70	-11.9
Apo A-I lipoprotein, rat	A24700	18	24	6	23	12	5.32	-13.7
proApo A-I lipoprotein, human	LPHUA1	16	30	6	21	17	5.35	-14.3
NADPH cytochrome P-450 reductase, rat	RDRTO4	41	60	21	38	36	5.07	-15.6
Retinol binding protein, human	VAHU	18	10	2	10	14	5.04	-16.9
Actin beta, rat	ATRTC	23	26	9	19	18	5.06	-17.2
Actin gamma, rat	ATRTC	20	29	9	19	18	5.07	-16.8
Apo A-I lipoprotein, human	LPHUA1	16	30	5	21	16	5.10	-17.5
Apo A-IV lipoprotein, human	LPHUA4	20	49	8	28	24	4.88	-19.7
Tubulin alpha, rat	UBRTA	27	37	13	19	21	4.66	-19.8
F1ATPase beta, bovine	PWBOB	25	36	9	22	22	4.80	-21.0
Tubulin beta, pig	UBPGB	26	36	10	15	22	4.49	-22.5
Protein disulphide isomerase (PDI), rat hepatic	ISRTSS	43	51	11	51	9	4.07	-25.0
Cytochrome b5, rat	CBRT5	10	15	6	10	4	4.59	-26.0
Apo C-II lipoprotein, human	LPHUC2	4	7	0	6	1	4.44	-30.5
Amino acid pI assumed in calulation:		3.9	4.1	6.0	10.8	12.5		

JOURNAL OF BIOLUMINESCENCE AND CHEMILUMINESCENCE VOL 4 59—78 (1989)

High Specific Activity Chemiluminescent and Fluorescent Markers: their Potential Application to High Sensitivity and 'Multi-analyte' Immunoassays

Roger Ekins*, Frederick Chu and Jacob Micallef
Department of Molecular Endocrinology, University College and Middlesex School of Medicine, University of London, Mortimer Street, London W1N 8AA, UK

The sensitivities of immunoassays relying on conventional radioisotopic labels (i.e. radioimmunoassay (RIA) and immunoradiometric assay (IRMA)) permit the measurement of analyte concentrations above *ca* 10^7 molecules/ml. This limitation primarily derives, in the case of 'competitive' or 'limited reagent' assays, from the 'manipulation errors arising in the system combined with the physicochemical characteristics of the particular antibody used; however, in the case of 'non-competitive' systems, the specific activity of the label may play a more important constraining role. It is theoretically demonstrable that the development of assay techniques yielding detection limits significantly lower than 10^7 molecules/ml depends on:

(1) the adoption of 'non-competitive' assays designs;
(2) the use of labels of higher specific activity than radioisotopes;
(3) highly efficient discrimination between the products of the immunological reactions involved.

Chemiluminescent and fluorescent substances are capable of yielding higher specific activities than commonly used radioisotopes when used as direct reagent labels in this context, and both thus provide a basis for the development of 'ultra-sensitive', non-competitive, immunoassay methodologies. Enzymes catalysing chemiluminescent reactions or yielding fluorescent reaction products can likewise be used as labels yielding high effective specific activities and hence enhanced assay sensitivities.

A particular advantage of fluorescent labels (albeit one not necessarily confined to them) lies in the possibility they offer of revealing immunological reactions localized in 'microspots' distributed on an inert solid support. This opens the way to the development of an entirely new generation of 'ambient analyte' microspot immunoassays permitting the simultaneous measurement of tens or even hundreds of different analytes in the same small sample, using (for example) laser scanning techniques. Early experience suggests that microspot assays with sensitivities surpassing that of isotopically based methodologies can readily be developed.

Keywords: Ultrasensitive immunoassay; fluorescent microspot immunoassay; confocal microscopy

*Author for correspondence.

0884-3996/89/030059-20$10.00

INTRODUCTION

Immunoassay methods relying on radioisotopic labels have played a major role in medicine and other biologically related fields (agriculture, veterinary science, the food and pharmaceutical industries, etc.) during the past two decades. Their importance has derived from the exploitation both of the 'structural specificity' characterizing antibody–antigen reactions and the 'detectability' of isotopically-labelled reagents, the latter permitting observation of the binding reactions between exceedingly small concentrations of the key reactants involved. The combination of these features has endowed radioimmunoassay methods with unique specificity and sensitivity characteristics, and accounts for their ubiquitous use throughout modern medicine and biology. However, in the past few years, interest has increasingly focused on so-called 'alternative', non-radioisotopic, immunoassay methods; such techniques are based on essentially identical analytical principles but differ in the markers used to label the particular immunoreactant (antibody or analyte) whose distribution between bound and free moieties (following the basic analytical reaction) constitutes the assay 'response'. The reasons for this interest may be grouped under four headings:

(1) Environmental; logistic; economic; practicality and convenience, etc. (i.e. 'non-scientific).
(2) The attainment of higher sensitivity.
(3) The development of 'immunosensors' and 'immunoprobes'.
(4) The development of 'multi-analyte' assay systems.

Our own reasons for developing non-isotopic techniques fall principally under headings (2) and (4), and this presentation will centre primarily on the concepts which underlie our immunoassay development strategy in these areas.

THE ATTAINMENT OF 'ULTRA-HIGH' IMMUNOASSAY SENSITIVITY

Though, as indicated above, the sensitivity of radioisotopically based immunoassay methods has constituted one of the principal foundations of their widespread use over the past 25 years, a fundamental reason for their replacement stems, paradoxically, from the current requirement to develop microanalytical techniques which are superior to them in this particular respect. Radioisotopic methods are, in practice, limited to the measurement of analyte concentrations above about 10^8–10^9 molecules/ml (i.e. approx 0.15–1.5 pmol/l)(Dakubu *et al.*, 1984). However, in certain fields (e.g. virology, tumour detection) there is a particular need to detect or measure molecular concentrations below this level. The factors which determine immunoassay sensitivity have been extensively discussed (Ekins *et al.*, 1968, 1970a; Ekins, 1978; Jackson *et al.*, 1983; Dakubu *et al.*, 1984; Ekins, 1985). Nevertheless, some of the underlying concepts are still frequently misunderstood and merit brief discussion in the present context.

The concept of sensitivity

One major source of past confusion has been disagreement regarding the concept of 'sensitivity' itself, many authors equating assay sensitivity with the slope of the dose–response curve (Yalow and Berson, 1970a, b; Berson and Yalow, 1973; see also Ekins *et al.*, 1970b, Tait, 1970). It is now widely agreed that the notion that a steeper dose–response curve implies greater sensitivity is erroneous. The invalidity of this belief is clearly revealed by the fact that the relative magnitudes of the responses yielded by two assay systems is dependent on the particular variable which is chosen to represent the response (see Fig. 1(a))(Ekins, 1976). For this and other reasons, it has long been recognized that the 'sensitivity' of an assay can only be satisfactorily represented by its lower limit of detection (Fig. 1(b)), and this concept is now embodied in all internationally agreed definitions of the term. An essentially identical definition is as the precision (i.e. standard deviation) of measurement of zero dose, since this quantity determines the least quantity distinguishable from zero and hence the assay detection limit. The sensitivity of an assay is thus represented by the zero-dose intercept of the 'precision profile' (Fig. 2(a)) when the latter is expressed in terms of standard deviation rather than of coefficient of variation (Ekins, 1983a). In short, the more sensitive of two assays is the one yielding greater precision of the zero dose estimate (Fig. 2(b)).

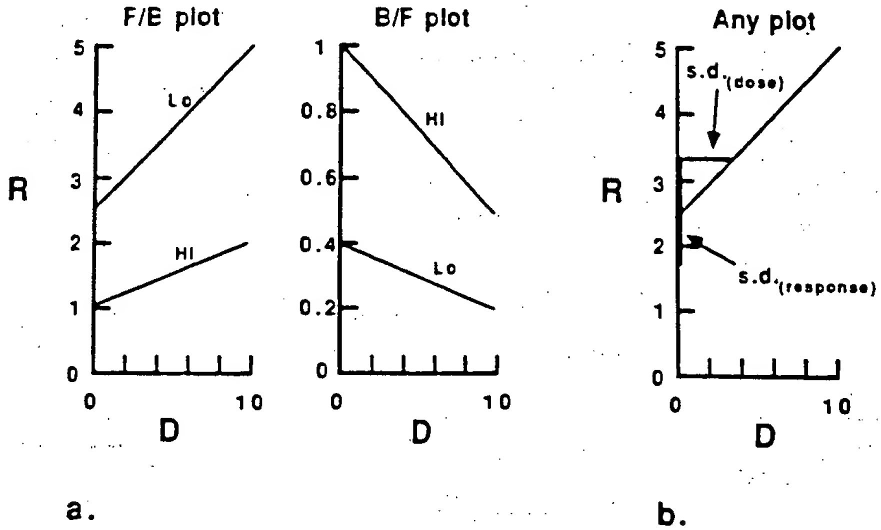


Figure 1. (a) Diagrammatic representation of conventional RIA dose–response curves for systems using high (hi) and low (lo) antibody concentrations plotted in terms of free-bound (F/B) and bound/free (B/F) labelled antigen. Note that the use of a lower amount of antibody yields a dose–response curve of greater slope in the F/B plot, but of lower slope in the B/F plot. It is impossible to decide, on the basis of the data shown in this figure, which concentration of antibody yields the assay system of higher sensitivity. (b) The sensitivity of an assay is essentially represented by the minimum detectable dose, i.e. the SD of the dose measurement ($SD_{(dose)}$) at zero dose. This is given by the SD of the response ($SD_{(response)}$) divided by the dose–response curve slope at zero dose (i.e. $((SD_{(response)}) \times dD/dR)_0$). This quantity is unaffected by the choice of the coordinate frame used to plot the dose–response curve. (Note: it is common to multiply $(SD_{(dose)})_0$ by an arbitrary factor to increase the confidence level attaching to the minimum detectable dose estimate, though, since no agreement exists regarding the value of this factor, this unnecessary step merely adds to confusion when the relative sensitivities of two assay procedures are compared.)

'Competitive' and 'non-competitive' ('limited reagent' and 'excess reagent') assays

A second important misconception in this area is the notion that immunoassays relying on the use of *labelled antibodies* (e.g. immunoradiometric assays, IRMA) are *ipso facto* more sensitive than those which rely on the use of *labelled 'analyte'* (e.g. radioimmunoassays, RIA); furthermore the grounds originally advanced for the claimed superiority of labelled antibody methods (Miles and Hales, 1968) were partially based on false concepts of sensitivity, and thus failed to identify the *true* reasons why certain assay designs are

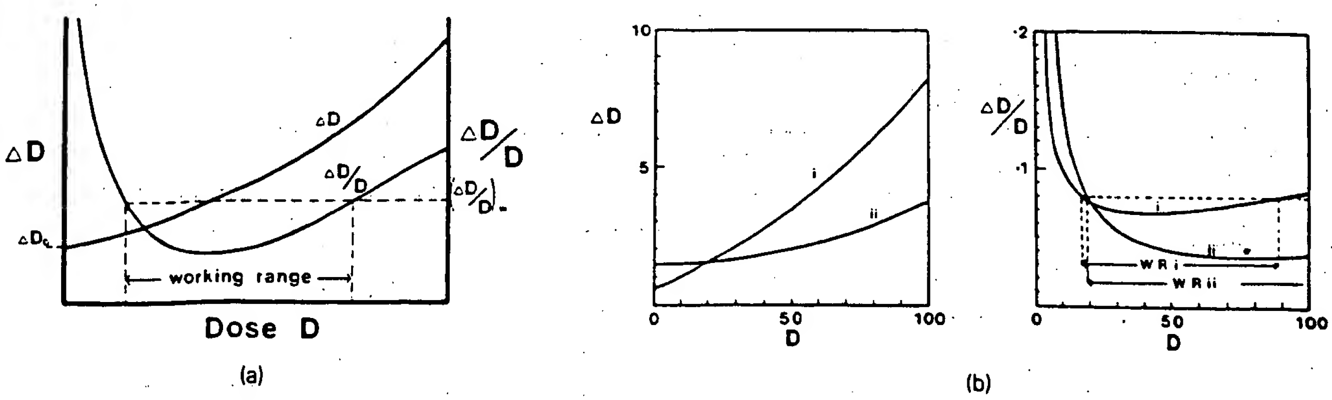


Figure 2. (a) The 'precision profile' of an assay portrays the error in the dose measurement as a function of dose. The error may be represented, *inter alia*, by the absolute error (ΔD; e.g. SD of D) or the relative error ($\Delta D/D$; e.g. CV of D). $(\Delta D)_0$, the error in the measurement of zero dose, represents the sensitivity of the assay. The working range may be defined as the range of dose values within which $\Delta D/D$ is less than an 'acceptable' value set by the investigator. (b) The more sensitive of the two assays (assay I) intercepts the ΔD axis at a lower value. However, assay II is more precise at higher values of dose, and has a wider working range.

potentially capable of yielding far higher sensitivity than others. This issue likewise merits clarification.

The purely pragmatic sub-classification of immunoassays into labelled antibody and labelled analyte methods diverts attention from a more fundamental divide in immunoassay methodology, which relates to the optimal concentration of antibody required in an assay system to maximize its sensitivity. In certain assay designs (which may be termed *'limited reagent'* or *'competitive'*) the optimal concentration tends to zero; conversely in others (which may be termed *'excess reagent'* or *'non-competitive'*) the concentration tends to infinity. It should be particularly emphasized that the optimal antibody concentration is essentially governed, not only by the physicochemical characteristics of the antibody–analyte binding reaction, but also by the errors incurred in measurement of the assay response. Were an assay system to be totally error-free, *no* antibody concentration would be optimal, and the distinction between competitive and non-competitive methodologies would thus not arise.

Though it is inappropriate in this presentation to discuss in detail the statistical and physicochemical theory underlying this fundamental divergence in immunoassay design (see Ekins *et al.*, 1968, 1970a; Jackson *et al.*, 1983), the reason for it can perhaps be more readily understood if the basic principles of immunoassay are portrayed in a somewhat different way from that in which they are usually presented. All immunoassays essentially depend upon measurement of the 'fractional occupancy' by analyte of antibody binding sites following reaction of analyte with antibody (see Fig. 3(a)). Those techniques which implicitly rely on measurement of residual, *unoccupied*, binding sites optimally necessitate the use of concentrations of antibody tending to zero, and may be termed 'competitive', conversely those in which *occupied* sites are directly measured necessitate use of high antibody concentrations and are termed 'non-competitive' (Fig. 3(b)). This emphasizes that the differences in assay design characterizing so-called competitive and non-competitive methods are essentially unrelated to which component (if any) of the reaction system is labelled. Indeed immunoassays in which *no label of any kind is involved* can, on identical grounds, be subdivided into those of 'limited reagent' (or 'competitive') and 'excess reagent' (or 'non-competititve') design. Thus the distinction between these two forms of immunoassay simply reflects differences in the way that fractional antibody occupancy is determined, and the fact that it is generally undesirable—for reasons of accuracy—to measure a *small* quantity by estimating the difference between two *large* quantities. When an immunoassay relies on the measurement of unoccupied antibody binding sites, the total amount of antibody used in the system must be small to minimize error in the resulting (indirect) estimate of occupied sites.

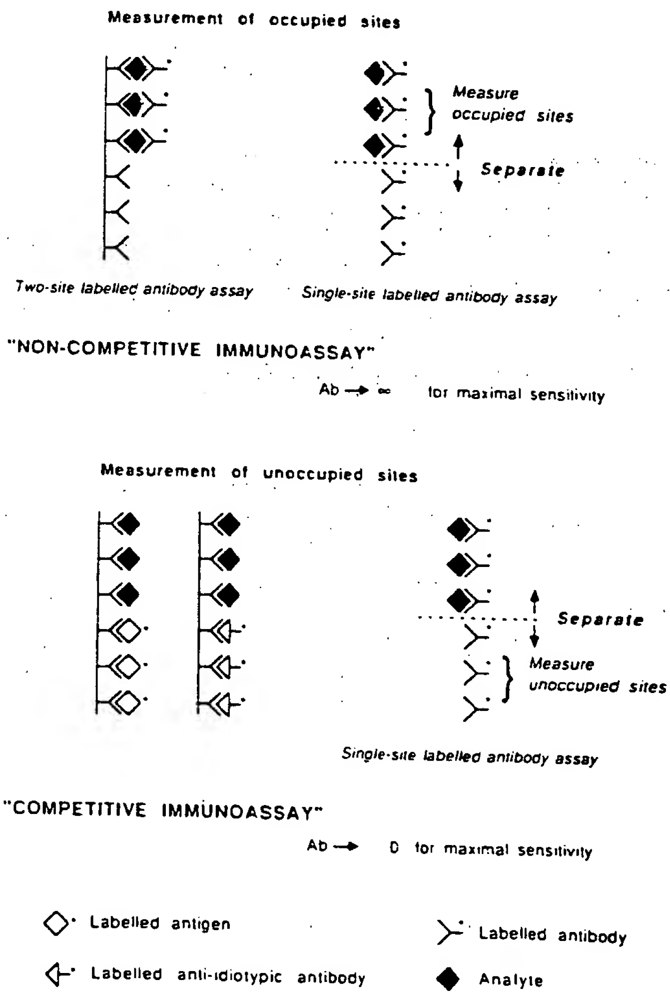


Figure 3. The distinction between 'non-competitive' (above) and 'competitive' immunoassays (below) reflects how antibody binding-site occupancy is measured. Labelled antibody methods are 'non-competitive' if occupied sites of the (labelled) antibody are measured, but are 'competitive' (below right) when *unoccupied* sites are measured. Labelled antigen (below left) or labelled anti-idiotypic antibody methods (below centre) rely on measurement of sites *unoccupied* by analyte, and are therefore invariably of 'competitive' design.

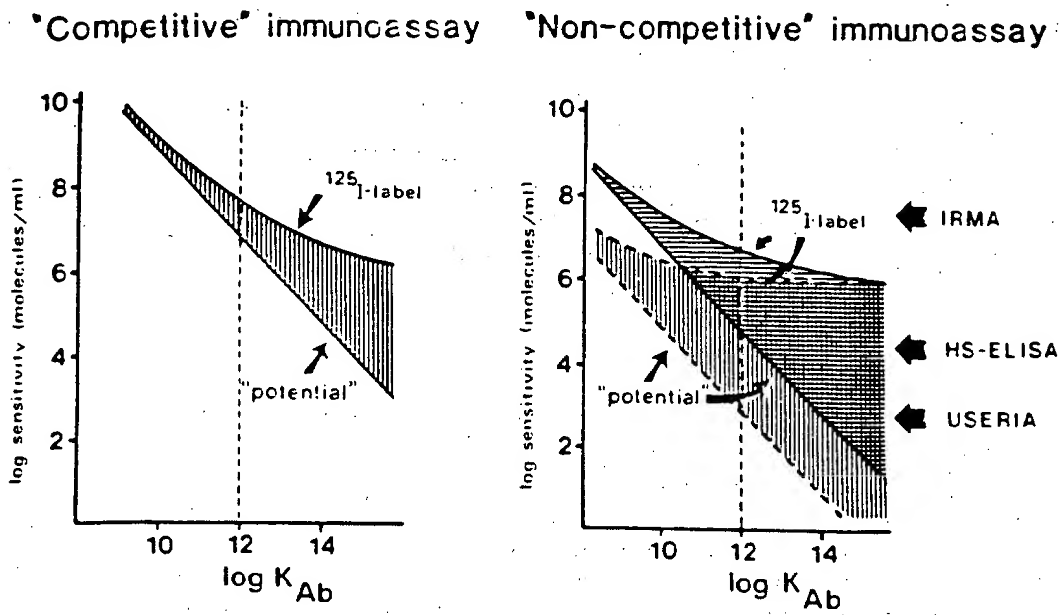


Figure 4. Curves showing the theoretically predicted relationship between antibody affinity and the sensitivities achievable using 'competitive' and 'non-competitive' assay strategies. The 'potential' sensitivity curves assume the use of infinite specific activity labels; the sensitivities achievable using ^{125}I-labelled antigen or antibody are also shown. Shaded areas indicate the sensitivity loss due to errors in measurement of the label. Curves relating to 'competitive' assays assume a 1% error in measurement of the response variable arising from 'experimental' errors (i.e. errors other than those inherent in label measurement *per se*). Non-competitive curves assume 'non-specific binding' of labelled antibody of 0.01% and 1% (lower and upper curves) respectively. Arrows indicate sensitivities claimed for typical non-competitive immunoassay methodologies.

Conversely, when occupied sites are measured *directly*, this particular constraint does not arise; indeed, considerable advantage often derives from using relatively large amounts of antibody in the system.

Sensitivity of 'competitive' and 'non-competitive' immunoassays

Competitive and non-competitive immunoassays differ significantly in many of their performance characteristics in consequence of the differences in optimal antibody concentration on which they rely. Most particularly they differ in their potential sensitivities. Figure 4. portrays the sensitivities predicted theoretically as a function of antibody binding affinity, making realistic assumptions regarding the experimental errors incurred in reagent manipulation, 'non-specific' binding of labelled antibody, etc., and assuming the use of optimal reagent concentrations (Ekins, 1985). Amongst other concepts illustrated in the figure is the much greater assay sensitivity *potentially* attainable (using an antibody of given affinity) by adoption of a non-competitive approach. In short, whereas the maximal sensitivity realistically achievable using a competitive design is in the order of 10^7 molecules/ml (using antibody of the highest affinity found in practice), a non-competitive method is capable of yielding sensitivities some orders of magnitude greater than this. However, Fig. 4 also demonstrates that, assuming the use of high affinity antibodies (i.e. $\sim 10^{11}$–10^{12} l/M), maximal sensitivities yielded by isotopically based techniques (whether relying on labelled antibody (IRMA) or labelled analyte (RIA), or whether of competitive or non-competitive design) are closely comparable, i.e. of the order of 10^7–10^8 molecules/ml.

This limitation is a manifestation of the fact that, in the case of the non-competitive methods, an important constraint on assay sensitivity is (under certain circumstances) the 'specific activity' of the label used. On the other hand, limitation of assay sensitivity due to the low specific activity of radioisotopic labels does *not* often arise, in practice, in the case of competitive assays, whose sensitivity is generally restricted by other factors (Ekins, 1985). The fundamental significance of this conclusion is that, only by the use of labels possessing specific activities higher than those of the commonly used radioisotopes *in assays of non-competitive design*, can current

sensitivity limits be breached. Conversely, use of a higher specific activity label in a *competitive* assay will usually have no significant effect on its sensitivity (assuming experimental errors incurred in reagent manipulation of the magnitude generally encountered in practice).

High specific activity non-isotopic labels

The term 'specific activity' is conventionally applied, in the case of radioisotopic labels, to denote the number of radioactive disintegrations per unit time per unit weight of the isotope or labelled compound. In the present context, use of the term is widened to signify 'detectable events' per unit time per unit weight of labelled material. Thus it can be used to indicate the rate of photon emission by a chemiluminescent or fluorescent label, or the rate of conversion of substrate molecules—by an enzyme label—to molecules of a detectable product. The importance of the concept derives from the fact that 'signal measurement error' (i.e. error in the measurement of the label *per se*) is a contributory factor in limiting assay sensitivity, and may—when other sensitivity-constraining factors are reduced—become dominant. Furthermore, when extending the sensitivities of immunoassay systems beyond their present limits, the numbers of molecules involved are low, and statistical errors incurred in counting individual 'detectable events', and the time required to count them, may assume a particular importance.

Table 1 compares the specific activities of potentially useful labels with that of ^{125}I. All are of relevance in the context of this volume since chemiluminescent and fluorescent labels can be used to label antibodies (or antigens) directly; alternatively, enzyme labels catalysing reactions yielding chemiluminescent signals or fluorescent products can be utilized.

Table 1. Relative specific activities of various isotopic and non-isotopic labels. Note that, though the specific activity of ^{125}I-labelled reagents does not, in practice, significantly limit the sensitivity of competitive assays (see Fig. 4), the lower specific activity of ^{3}H may severely restrict the sensitivity of competitive assays (e.g. of steroid hormones) which rely on the use of this particular radioisotope

Specific Activities	
^{125}I:	1 detectable event/sec/7.5 × 10^6 labelled molecules.
^{3}H:	1 detectable event/sec/5.6 × 10^8 labelled molecules.
Enzymes:	Determined by enzyme 'amplification factor' and detectability of reaction product.
Chemiluminescent labels	1 detectable event/labelled molecule.
Fluorescent labels:	Many detectable events/labelled molecule.

The importance of background in non-competitive immunoassays

A second important factor governing the sensitivity of non-competitive labelled-antibody immunoassays is the 'background' or 'blank' signal emitted in the absence of analyte, since error in the measurement of this signal is clearly a major determinant of the error in measurement of zero dose. Amongst contributors to the background signal are the 'noise' of the measuring instrument itself, 'ambient' signal generators (such as, in 'sandwich' immunoassays, solid 'capture-antibody' supports or, in the case of radioisotopic methods, cosmic ray and other extraneous radiation sources) and 'non-specifically bound' labelled antibody. Minimization of each of these components is essential for maximal sensitivity: mere arithmetic subtraction of background is of absolutely no benefit in this context.

Non-specific binding of antibody is of particular interest, since the magnitude of this contribution is dependent, *inter alia*, on the amount of labelled antibody used in the system, and the duration of its exposure to analyte. Thus increasing the amount of labelled antibody increases the amount of such antibody bound to analyte; however, it may also increase the non-specifically bound moiety to a greater proportional extent, and thus cause a net reduction in sensitivity. This effect underlies the loss in sensitivity at higher antibody concentrations depicted in Fig. 5 (reproduced from Jackson *et al.*, 1983). This phenomenon also underlies the relationship between sensitivity and the affinity constant of the labelled antibody depicted in Fig. 4. The possession by labelled antibody of a high affinity constant implies that a

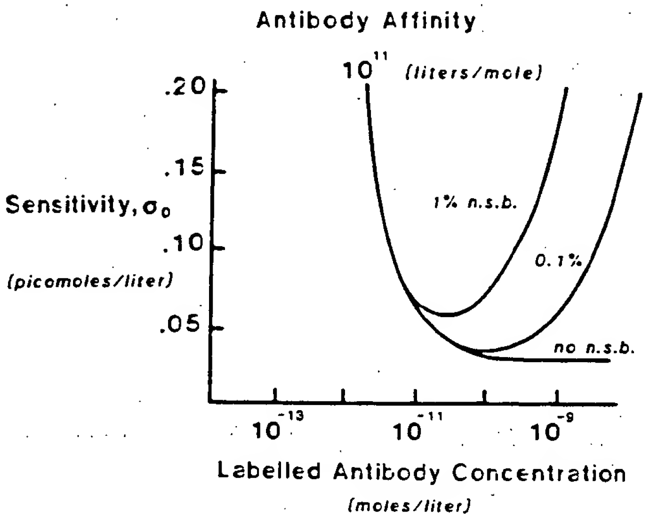


Figure 5. Assay sensitivity (represented by the standard deviation of the zero dose measurement, σ_0), plotted as a function of the concentration of labelled antibody (of affinity 10^{11} L/M) used in the assay, assuming different levels of non-specific binding of labelled antibody. (Note: an irreducible instrument background has been assumed in the computations represented; this limits the ultimate sensitivity attainable, regardless of the concentration of antibody used.)

lower concentration is required to yield the same level of analyte binding, albeit with reduced non-specific binding, thus increasing assay sensitivity

In summary, the high sensitivity of non-competitive labelled antibody methods derives essentially from their permitted use of optimal concentrations of antibody which (provided non-specific binding of labelled antibody is low) are generally considerably greater than in competitive methods, *not* from the fact that the antibody is labelled. Labelled antibody methods generally *fall* in sensitivity as the concentration of antibody is reduced towards zero, ultimately yielding a sensitivity theoretically identical to that of competitive methods (Rodbard and Weiss, 1973). (Paradoxically, early exponents of labelled antibody methods, whilst claiming them to be of higher sensitivity, also concluded that their sensitivity was *increased* by reduction in the amount of labelled antibody used (Woodhead *et al.*, 1971). This incorrect conclusion—based on observation of effects on the slope of the dose-response curve—exemplifies the many fallacies encountered in the immunoassay field stemming from confusion regarding the concept of sensitivity discussed above.) Finally it should be emphasized that maximization of the sensitivity of a non-competitive immunoassay generally implies the selection of reagent concentrations and other experimental conditions such that the [analyte signal/background] ratio (i.e. s/b) is maximized. However, this simple relationship disregards statistical considerations which arise when the numbers of detectable events are very low, and a more appropriate objective may, under these circumstances, be maximization of the ratio s^2/b (Loevinger and Berman, 1951).

Other performance characteristics of competitive and non-competitive immunoassays

Non-competitive designs also display a number of other advantages deriving from the relatively high antibody concentrations on which they generally rely. These include increased reaction speeds (and hence shorter incubation times), decreased vulnerability to certain environmental effects (which cause variations in binding affinity between antibody and analyte), reduced sensitivity-dependence on high antibody binding affinity, etc.

Nevertheless a price has to be paid for these benefits; this includes the greater tendency of a large amount of antibody to bind molecules differing from, but with structural resemblance to, the analyte itself, implying a loss of assay *specificity*. This effect generally necessitates the use, whenever possible, of an 'immunoextraction' procedure using a second 'capture' antibody (usually directed against a different binding site, or 'epitope') as shown in Fig. 3(b). This technique—the 'sandwich' or 'two-site' immunoassay (Wide, 1971)—thus potentially combines the twin virtues of ultra-high sensitivity and specificity (together with short reaction time), features of crucial importance in many diagnostic situations (for example, in the detection of AIDS viral antigens). (Note, however, that the loss of specificity inherent in non-competitive assay designs implies that they are less readily applicable to the measurement of analytes of small molecular size, which cannot be simultaneously bound by two different antibodies directed against different antigenic sites on the molecule. Such analytes are generally more appropriately measured using 'competitive' assay methods.)

Development of ultra-sensitive immunoassay methodologies

The perception that the development of 'ultra-sensitive' immunoassay systems (i.e. systems surpassing conventional RIA methods in sensitivity) depends on (a) reliance on 'excess reagent' or 'non-competitive' assay designs; (b) the use of non-isotopic labels displaying higher specific activities than commonly used radioisotopes; (c) the development of efficient separation systems (ensuring minimization of non-specific antibody binding, and hence of signal 'backgrounds'), and (d) dual or multi-antibody analyte-recognition systems (exemplified by 'sandwich' or two-site assays) to maintain/increase assay specificity, has formed the basis of our own laboratory's immunoassay development since the early to mid-1970s (Ekins, 1978). This led us, *inter alia*, to an immediate recognition (Ekins, 1979, 1980) of the importance of the *in vitro* techniques of monoclonal antibody production pioneered by Köhler and Milstein (1975), which are currently the subject of bitter patent disputes in the USA (Ezzell, 1986, 1987a,b); and which may be expected in Europe.

Meanwhile, of the candidate labels for use in this context, both chemiluminescent and fluorescent labels offer many attractions. The development of stable, highly chemiluminescent, acridinium esters by McCapra and his colleagues (McCapra *et al.*, 1977) has subsequently been exploited by Weeks *et al* (1983, 1984) and, more recently, by several commercial kit manufacturers; other workers have used more conventional chemiluminescent compounds to label immunoassay reagents (see, for example, Kohen *et al.*, 1984, 1985; Barnard *et al.*, 1985). Yet others have relied on enzyme labels to catalyse chemiluminogenic (Whitehead *et al.*, 1983) and fluorogenic (Shalev *et al.*, 1980) reactions as indicated above. Detailed description of these various methodologies is presented by others in this volume and need not be duplicated here.

Common to all the 'ultra-sensitive' immunoassay methodologies relying on such alternative labels is their dependence on a non-competitive, labelled antibody, assay strategy whenever appropriate; however, for the reasons indicated above, *competitive* methods continue to be generally employed for the measurement of analytes of small molecular size (e.g. therapeutic drugs, steroid and thyroid hormones, etc.). Nevertheless, the convenience (from a manufacturing viewpoint, and for other technical reasons) of relying on standard labelling procedures has meant that, even in these cases, labelled antibody techniques are increasingly preferred. Though the commercial kits based on these various labels differ to a minor extent in sensitivity, specificity, convenience, etc., such differences are at least partially attributable to differences in the physicochemical characteristics of the antibodies used in the kits, and to other 'immunological' factors unconnected with the particular nature of the label *per se*.

Despite the obvious attractions of chemiluminescent techniques in an immunoassay context, the use of fluorescent labels combined with sophisticated time-resolution techniques for their detection (a concept arising from discussions with J. F. Tait in 1970) appeared to us (in the mid-1970s) to offer more exciting long-term possibilities for a number of reasons. These naturally included attainment of the enhanced specific activities and high signal to background ratios required for ultra-sensitive immunoassay as indicated above. However, more importantly, fluorescence techniques also appeared to provide a simple route to the development of 'multi-analyte' assay systems of the kind described below.

In pursuance of this strategy, we began collaboration with LKB/Wallac, *ca* 1976–77, in the development of the instrumentation and technology required to develop such methods. Fortunately a group of fluorescent substances generally known as the lanthanide chelates (including, in particular, the chelates of europium, samarium and terbium facilitate such development, possessing prolonged fluorescence decay times (~10–1000 μs), large Stokes shift (~300 nm) and other desirable physical characteristics which permit the construction of relatively cheap instrumentation for their measurement (Marshall *et al.*, 1981; Hemmilä *et al.*, 1983). The fluorescent properties of the lanthanide chelates may be compared with those of a conventional fluorophor such as fluorescein which is characterized by a much smaller Stokes shift (~28 nm), and a fluorescent decay time and emission spectrum which imply that it is less readily distinguished from fluorescent substances present in blood (such as bilirubin) or in plastic sample holders. The unique fluorescence characteristics of the lanthanide chelates thus permit them to be

measured in the presence of a fluorescence background (deriving from extraneous sources) which, in practice, approaches zero. Fig. 6 illustrates the basic concepts involved in pulsed-light, time-resolved, fluorescence measurement, which form the basis of the DELFIA immunoassay system currently marketed by LKB/Wallac.

Though it is inappropriate to pursue this subject in greater detail, attention should also be drawn to the possibilities offered by phase-resolved fluorimetry. This permits separate identification of fluorophores differing in fluorescence lifetime by their exposure to a sinusoidally modulated exciting light source, and observation of their demodulated, phase-shifted, light emission (McGown and Bright, 1984). This technique offers the possibility both of the development of homogeneous assays (relying on a difference in fluorescence decay time of bound and free forms of the fluorescent-labelled molecule), and of discriminating between two labelled antibodies in the context of multi-analyte 'ratiometric' immunoassay as discussed below.

'AMEIENT ANALYTE' IMMUNOASSAY

Before proceeding to a discussion of the development of multi-analyte assays, another important concept, termed 'ambient analyte immunoassay' (Ekins, 1983b), must first be examined. This term is intended to describe a type of immunoassay system which, unlike unconventional methods, measures the analyte *concentration* in the medium to which an antibody is exposed, being essentially independent both of sample volume, and of the amount of antibody present. This concept is illustrated in Fig. 7, and relies on the physicochemically-based proposition that, when a 'vanishingly small' amount of antibody (preferably, but not essentially, coupled to a solid support) is exposed to an analyte-containing medium, the resulting (fractional) occupancy of antibody binding sites solely reflects the ambient analyte concentration. Clearly the binding by antibody of analyte results in a depletion of the amount of analyte in the surrounding medium, but provided the proportion so bound is small (i.e. less than, for example, 1% of the total), such disturbance can be ignored. (This effect is closely analogous to that caused by the introduction of a thermometer into a medium possessing a much larger thermal capacity; the temperature disturbance caused by the thermometer itself is negligible and can, in these circumstances, be disregarded.)

The principles of ambient analyte assay derive from the recognition that *all* immunoassays essentially depend upon measurement of the 'fractional occupancy' by analyte of antibody binding sites following reaction of analyte with antibody as discussed above (Figs 3. (a) and (b)). The fractional occupancy of ('monospecific' or 'monoclonal') antibody binding sites in the presence of varying analyte concentrations, plot-

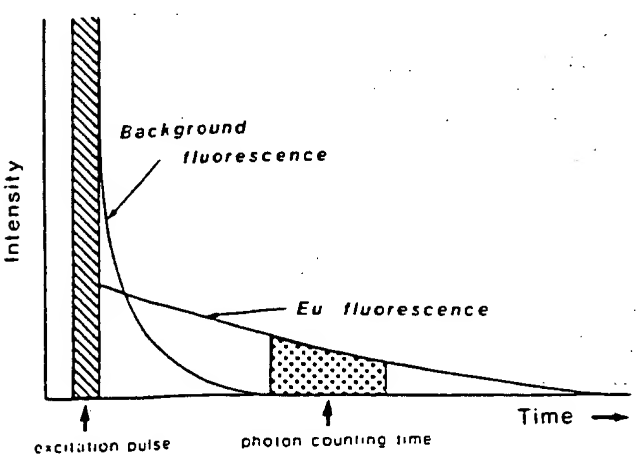
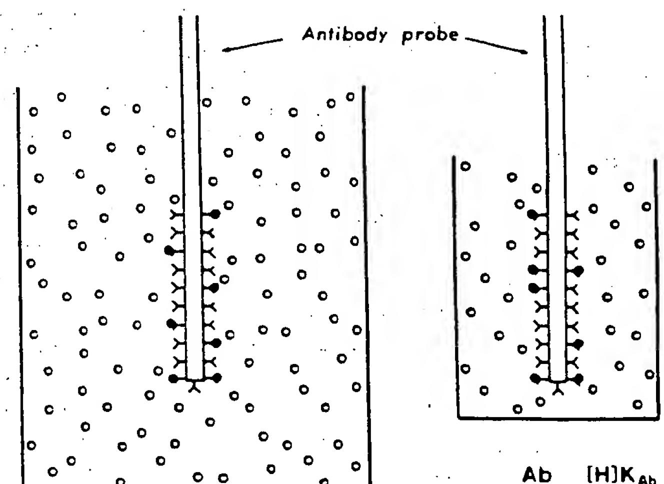


Figure 6. Basic principles of pulse-light, time resolved fluorescence. Fluorescence emitted by the fluorophor (typically a europium chelate) is distinguished from background fluorescence, which decays more rapidly.



$$\frac{Ab}{Ab_0} = \frac{[H]K_{Ab}}{1-[H]K_{Ab}}$$

Figure 7. Basic principle of 'ambient analyte' immunoassay (AAI). The fractional occupancy (*F*) of a vanishingly small amount of antibody (of affinity *K*) is determined by the analyte concentration in the medium ([An]).

ted against antibody concentration, is portrayed in Fig. 8. The fraction of analyte bound is also plotted in this figure. (Note: for the sake of generality, all concentrations in this figure are expressed in terms of $1/K$, where K is the affinity constant of the antibody. For example, if $K = 10^{11}$ L/M, a concentration of $0.1 \times 1/K$ represents 0.1×10^{-11} M/L, or $0.1 \times 10^{-11} \times 10^{-3} \times 6.02 \times 10^{23} = 6.02 \times 10^{8}$ molecules/ml.)

It should be particularly noted that, at antibody concentrations of less than *ca* $0.01 \times 1/K$ antibody fractional occupancy is essentially dependent solely on the analyte concentration in the medium, and is independent of variations in antibody concentration. This reflects the fact that this concentration of antibody binds less than approximately 1% of the analyte in the medium, irrespective of its concentration. This implies, for example, that the introduction of 10, 100, or 1000 antibody molecules into a medium containing billions of analyte molecules will result, in each case, in virtually identical fractional antibody binding-site occupancy, the upper limit of antibody concentration being determined by the antibody affinity constant. (An antibody concentration of $0.01 \times 1/K$ is a hundred-fold less than that ($1 \times 1/K$) necessary to bind 50% of a 'trace' amount of analyte (see Fig. 8), claimed by Berson and Yalow (1973) as maximizing assay 'sensitivity' (i.e. the slope of the dose–response curve when expressed in terms of bound/free labelled analyte). This false conclusion has subsequently become incorporated into the mythology of radioimmunoassay design which, regrettably, a majority of kit manufacturers continue to accept.)

The ambient analyte assay concept was originally exploited in the original development of what has come to be known as 'two-step' free hormone immunoassay (Ekins *et al.*, 1980), but it is clear that it is of far wider application, and can, in particular, be utilized in the construction of immunosensors and immunoprobes. One such example is a probe for the measurement of salivary steroids that is currently being developed in our laboratory. Comprising a small antibody-coated plastic 'dipstick' comparable in size and shape to a clinical thermometer, this device is intended to permit the measurement of salivary steroid levels without requiring the collection of saliva. However, the concept also underlies our approach to multi-analyte immunoassay, also under development in our laboratory.

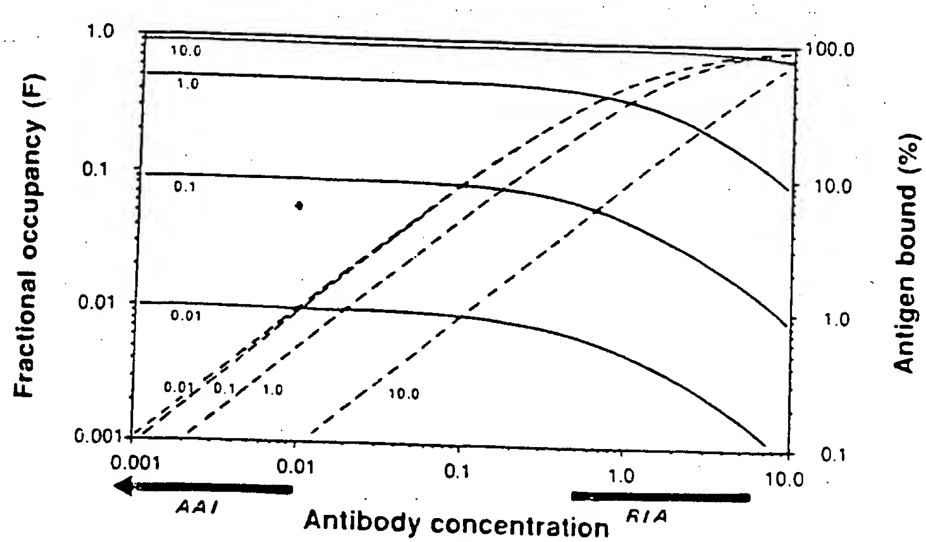


Figure 8. Fractional antibody binding-site occupancy (F) plotted as a function of antibody binding-site concentration for different values of analyte (antigen) concentration [An]. The percentage binding of analyte to antibody (b) is also shown. All concentrations are expressed in units of $1/K$. Note that for antibody concentrations of less than $0.01/K$ (approximately), percentage binding of analyte is <1%, and fractional binding-site occupancy is essentially unaffected by variations in antibody concentration extending over several orders of magnitude, being governed solely by [An]. Note that radioimmunoassays and other 'competitive' immunoassays are commonly designed using antibody concentrations approximately $0.5/K$–$1/K$ or above implying $b_0 > 30\%$), in accordance with the precepts of Berson and Yalow (e.g. Berson and Yalow, 1973).

MULTI-ANALYTE 'RATIOMETRIC' IMMUNOASSAY SYSTEMS

The concepts relating to ambient analyte immunoassay and assay sensitivity outlined above are both exploited in our present development of a random access, multi-analyte, immunoassay technology capable of measuring, in the same small sample, virtually any number of individual analytes from selected analyte 'menus' (e.g. a hormone menu, viral antigen menu, an allergen menu, etc.). Many examples of a need to measure a multiplicity of different analytes in the same sample exist in medical diagnosis, for example, in the routine diagnosis of thyroid disease, where it is frequently necessary to measure a number of different hormones and thyroid-related proteins. At present, clinicians frequently experience difficulty in deciding on the best sequence of tests to arrive at a correct diagnosis. Such problems would be overcome were all relevant analytes measurable at a cost comparable to the cost of measurement of a single substance. Our own immediate objective is the development of a technology permitting the measurement of complete 'hormone profiles' using a single small blood sample. However, the need for 'multi-analyte', or 'random access' measurement is not confined to medical diagnosis: it also arises, for example, in the pharmaceutical industry (where there exists a requirement to ensure the purity of protein drugs synthesized by recombinant DNA techniques), in the food industry and elsewhere. Though still at an early stage, our approach to the achievement of this objective can be briefly indicated.

Multi-analyte assay: general principles

As discussed above, the notion of ambient analyte assay simultaneously introduces two extremely important and novel concepts: (a) that an estimate of analyte concentration can be based upon the use of an infinitesimal amount of 'sampling' antibody, and (b) that such an estimate derives from a direct measurement of fractional antibody occupancy by analyte, irrespective of the exact amount of antibody used. It should be emphasized that the latter proposition is valid only in the context of ambient analyte assay, and is *not* true in current conventional immunoassay systems (in which fractional antibody occupancy depends both upon the amount of antibody in the system, and sample volume—see Fig. 8). In short, exposure of a small number of antibody molecules (in the form, for example, of a 'microspot' located on a solid support) to an analyte-containing fluid results in occupancy of antibody binding sites in the microspot reflecting the analyte concentration in the medium. Following such exposure, the antibody-bearing probe may be removed and exposed to a 'developing' solution containing a high concentration of an appropriate second antibody directed against either a second epitope on the analyte molecule if this is large (i.e. the occupied site), or against unoccupied antibody binding sites in the case of small analyte molecules (see Fig. 3(b)). (Note: an antibody simulating antigen, and reacting with unoccupied binding sites, is described as a 'mirror-image anti-idiotypic antibody'; the use of such an antibody instead of labelled antigen is convenient but not essential, and is suggested here merely to simplify illustration of the basic concepts involved.)

Subsequently, an estimate of binding-site occupancy of the 'sampling' (solid phase) antibody located in the microspot may be derived by measurement of the ratio of signals emitted by the two antibodies forming the dual-antibody 'couplets'. This can be conveniently achieved by labelling the 'sampling' and 'developing' antibodies with different labels, for example, a pair of radioactive, enzyme or chemiluminescent markers. Fluorescent labels are nevertheless particularly useful in this context because, by the use of optical scanning techniques, they permit arrays of different antibody 'microspots' distributed over a surface, each directed against a different analyte, to be individually examined, thus enabling multiple assays to be simultaneously carried out on the same small sample. Fig. 9 illustrates these basic ideas, and Fig. 10 such an array.

Microspot immunoassay sensitivity: theoretical considerations

The notion that it is, in principle, possible to measure an analyte concentration using a microspot of antibody comprising a number of antibody molecules in the range *ca* 10^1–10^6 is likely, at first sight, to appear surprising, and may, indeed, provoke scepticism regarding the assay sensitivities potentially attainable using this approach. Clearly a number of factors, such as the sensitivity

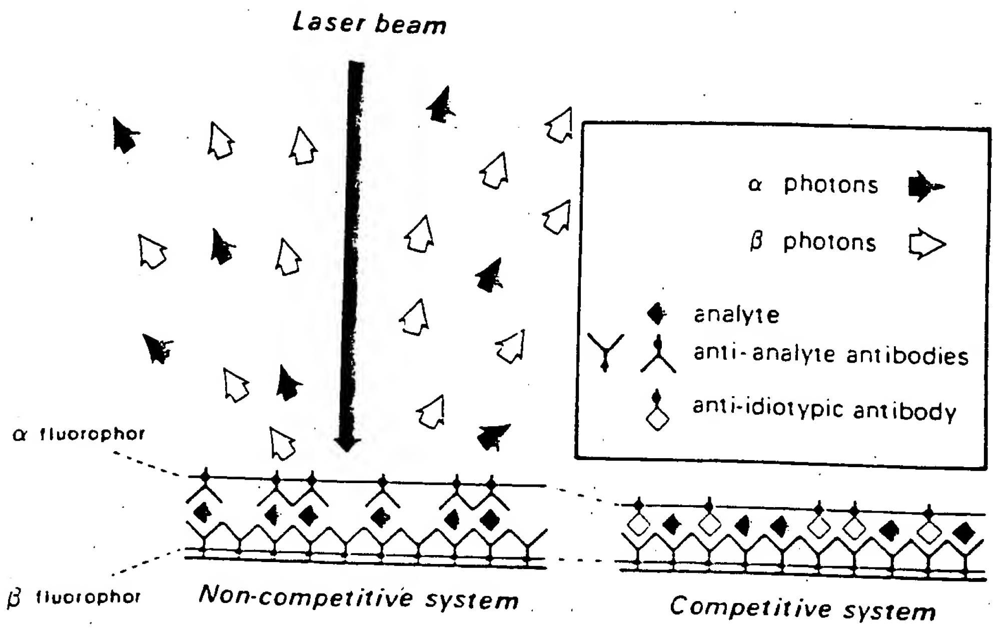


Figure 9. Basic principle of dual-label, ambient-analyte, immunoassay relying on fluorescent labelled antibodies. The ratio of α and β fluorescent photons emitted reflects the value of F (see Figs 5 and 6) and is solely dependent on the analyte concentration to which the probe has been exposed. It is unaffected by the amount or distribution of antibody coated (as a monomolecular layer) on the probe surface.

of the signal measuring equipment, the density of antibody molecules on the surface of the solid support, etc., are likely to play a part in determining final assay sensitivity. Such factors are, in turn, dependent on the efficiency with which the particular labels used can be detected, the adsorption properties of antibody supports,

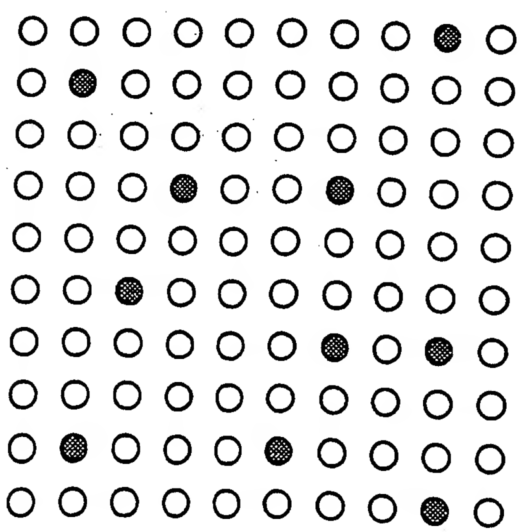


Figure 10. 'Multi-analyte' antibody array. Each antibody 'microspot' represents a 'vanishingly small' amount of antibody directed against an individual analyte.

etc. Though these are obviously variable, reasonable estimates can be made of the order of sensitivities likely to be achieved on the basis of some simple theoretical calculations. To clarify the following discussion, it is assumed that 'sensing' antibody can be uniformly and consistently coated on a solid matrix at a standard density, implying that only the 'developing' antibody need be labelled and measured in order to ascertain fractional occupancy of sensing antibody binding sites.

Fig. 11 illustrates the surface of an antibody microspot, of surface area $A(\mu m^2)$, and (uniformly) coated with antibody of affinity K(L/M) in a monomolecular layer of density D(molecules/μm^2). Let us assume that the spot is exposed to an analyte-containing medium of volume v(ml), and containing an analyte concentration C molecules/ml. The molecular concentration of antibody in the system is thus given by AD/v. (Note: the fact that antibody is situated on the surface of a solid support, and not evenly distributed throughout the medium, does not affect the extent of analyte binding at thermodynamic equilibrium, assuming that antibody binding sites are not impeded in their reactions and have not been damaged during the coating process.)

Meanwhile, fractional occupancy (F) of antibody binding sites by analyte (at equilibrium) is

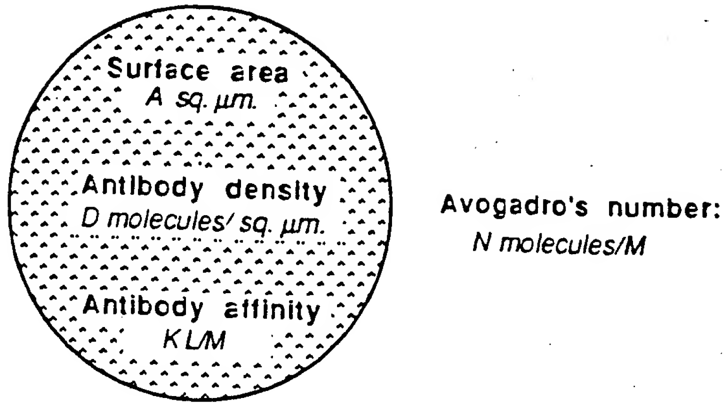


Figure 11. Microspot ambient-analyte immunoassay. The microspot shown is assumed to be uniformly coated with antibody, though if the dual-labelled antibody 'ratiometric' approach shown in Fig. 9 is adopted, uniform coating is not essential. The minimum fluid volume for ambient analyte assay conditions to prevail (enabling adoption of the ratiometric approach) is shown. Minimum test sample volume (M/S): $A \times D \times K \times 10^5/N$

given by the equation:

$$F^2 - F(1/q + p/q + 1) + p/q = 0 \quad (1)$$

where p = analyte concentration, q = antibody concentration (both expressed in units of $1/K$).

Thus, for antibody binding site concentrations $\rightarrow 0$ (i.e. $q < 0.01$), $F \approx p/(1 + p)$; (see Fig. 8).

Likewise, the fraction of analyte bound by antibody (f) at equilibrium is given by the equation:

$$f^2 - f(1/p + q/p + 1) + q/p = 0 \quad (2)$$

Thus, for analyte concentration $\rightarrow 0$ (i.e. $p < 0.01$), $f \approx q/(1 + q)$; (see Fig. 8). Furthermore, when $q < 0.01$, and when $p \geqslant 0$, $f < 0.01$.

Expressed in units of $1/K$; the concentration (q) in the assay of 'sensing' antibody situated on the microspot is given by $DAK/(\nu \times 6 \times 10^{20})$, (since Avogadro's constant, expressed as the number of molecules/mmol, is 6×10^{20} (approximately)). The fraction of an analyte concentration $\rightarrow 0$ which will be bound to the spot is therefore $DAK/(\nu \times 6 \times 10^{20} + DAK)$, implying that the number of analyte molecules bound to the spot is given by $\nu CDAK/(\nu \times 6 \times 10^{20} + DAK)$.

Case 1: sandwich (two-site) assay. Following incubation of sample with antibody, we assume the sample is removed, and the microspot then exposed to a volume V(ml) of a solution of a second, labelled, 'developing' antibody of affinity K^* (L/M) at a concentration given by Q (expressed in units of $1/K^*$).

The fraction of analyte bound by labelled antibody (F^*) at equilibrium is given by the equation:

$$F^{*2} - F^*(1/P + Q/P + 1) + Q/P = 0 \quad (3)$$

where P represents the analyte concentration in the developing-antibody solution, expressed in units of $1/K^*$, i.e. $\nu CDAKK^*/[(\nu \times 6 \times 10^{20} + DAK)V \times 6 \times 10^{20}]$.

Assuming $P < 0.01$, $F^* \approx Q/(1 + Q)$. (For example, if $Q = 1$, the fraction of analyte molecules bound by labelled antibody = 0.5 approximately). Thus, since the number of analyte molecules bound to the spot is given by $\nu CDAK/(\nu \times 6 \times 10^{20} + DAK)$, the number of analyte molecules labelled by the second, developing, antibody is given by $\nu CDAKQ/[(\nu \times 6 \times 10^{20} + DAK)(1 + Q)]$, and the surface density of such molecules is given by $\nu CDKQ/[(\nu \times 6 \times 10^{20} + DAK)(1 + Q)]$. Moreover, assuming that $DAK \ll \nu \times 6 \times 10^{20}$ (i.e. that the amount of antibody in the system is such that 'ambient assay' conditions prevail, then the surface density (D^*) of developing-antibody molecules = $CDKQ/[(6 \times 10^{20})(1 + Q)]$ approximately. It should be noted that D^* is independent of both ν and V, also that the ratio $D^*/D = C \times KQ/[(6 \times 10^{20})(1 + Q)] = C \times$ constant.

If the minimum detectable surface density of developing-antibody molecules (i.e. σ_{D0}^*, the standard deviation of the measurement of D^* when $C = 0$) is given by $D^*_{\min}$ (molecules/μm^2) and $C_{\min}$ represents the minimum detectable analyte concentration in the test sample, then,

disregarding non-specific binding of developing antibody within the microspot area,

$$C_{min} = D^*_{min} \times [(6 \times 10^{20})(1 + Q)]/DKQ \quad (4)$$

For example, if $Q = 1$, $D = 10^5$ molecules/μm^2, $K = 10^{11}$ L/M and $D^*_{min} = 20$ molecules/μm^2, then $C_{min} = 2.4 \times 10^6$ molecules/ml $= 10^{-15}$ M/L. It should be noted, in this example, the fractional occupancy of the sensing antibody binding sites by the minimum detectable analyte concentration is 0.04%.

Case 2: anti-idiotypic antibody ('competitive') assay. In this case, we assume that, following removal of the sample, the microspot is exposed to a volume V(ml) of a solution of (for example) a second, labelled, anti-idiotypic antibody reacting with *unoccupied* sites on the sensing antibody. Using similar reasoning as above, we may likewise assume that the fraction of such sites which become occupied by the anti-idiotypic 'developing' antibody is given by $Q/(1 + Q)$, where Q is the developing-antibody concentration. However, the minimum detectable surface density of anti-idiotypic antibody is not, in a competitive design, the critical determinant of assay sensitivity; this parameter is essentially governed by the precision of the density measurement.

From Eq. (1), the fraction of sites *unoccupied* by analyte $= 1/(1 + p)$, and the fraction occupied by anti-idiotypic antibody $= Q/(1 + p)(1 + Q)$. Thus, if the CV in the measurement of anti-idiotypic antibody is ε, the standard deviation is $\varepsilon Q/(1 + p)(1 + Q)$. This term also represents the SD in the estimate of the fraction of sites *occupied* by analyte. Since the total number of antibody binding sites in the spot is DA, the SD in the estimate of occupied sites as $p \rightarrow 0$ (i.e. σD^*_0) approximates $\varepsilon DAQ/(1 + Q)$; the SD in the occupied site surface-density estimate is thus $\varepsilon DQ/(1 + Q)$. But the SD in the measurement of fractional binding-site occupancy when $p \rightarrow 0$ defines D_{min}, and hence the minimum detectable analyte concentration in the test sample as indicated in Eq (4).

Thus

$$C_{min} = D_{min} \times [(6 \times 10^{20})(1 + Q)]/DKQ \quad (5)$$

$$= \varepsilon DQ/(1 + Q) \pm [(6 \times 10^{20})(1 + Q)]/DKQ \quad (6)$$

$$= \varepsilon/K \times (6 \times 10^{20}) \quad (7)$$

For example, if values of $Q = 1$, $D = 10^5$ molecules/μm^2, and $K = 10^{11}$ L/M are assumed as in the non-competitive example considered above, and the CV in the measurement of anti-idiotypic antibody density in the microspot is 1% (i.e. $\varepsilon = 0.01$), then $D_{min} = 500$ molecules/μm^2, and $C_{min} = 6 \times 10^7$ molecules/ml $= 10^{-13}$ M/L. Fractional occupancy of the sensing antibody binding sites by the minimum detectable analyte concentration is, in this example, 1%. It should be noted that the sensitivity limit of ε/K (expressed in molar terms) is identical to that previously established for conventional 'competitive' assays (Ekins and Newman, 1970), and which underlies the predictions represented in Fig. 4.

Such considerations appear to suggest (a) that microspot assay sensitivities superior to those obtainable by conventional radioisotopically based immunoassays are achievable, and (b) that sensitivities yielded by non-competitive microspot assays are likely to be considerably greater than those of corresponding competitive microspot assays. It must be emphasized, however, that, though such predictions are likely to prove correct, assumptions regarding the performance of the labels and signal-measuring instrument used are incorporated in the simple theoretical analysis discussed above. Such factors are clearly of importance in determining overall microspot immunoassay performance.

Practical implementation

The concepts discussed above are clearly exploitable using a variety of antibody labels, including chemiluminescent labels; however, our preliminary studies have been based on the use of conventional fluorophores, since the technology of simultaneous measurement of dual fluorescence from small areas is already well established. Because this volume centres on chemiluminescence, we shall provide only a brief indication of our initial experimental work in this area, which is currently based on the use of commercially available confocal microscopes.

Instrumentation: the laser scanning confocal microscope. In laser scanning confocal fluoresc-

ence microscopy, a small area of the specimen is illuminated by a focused laser beam; the fluorescence photons emanating solely from this area are, in turn, focused onto a photon detector. Both the intensity of illumination and the efficiency of light collection diminish rapidly with distance from the focal plane (Fig. 12). At the 'confocal' point, the projection of the illumination pinhole and the back-projection of the detector pinhole coincide. Such systems contrast with conventional epi-fluorescence methods, where the specimen is exposed to an essentially uniform flux of illumination (White *et al.*, 1987).

Sensitivity of current instruments. Typically, fluorescence photons emanating from the laser-illuminated area are detected by a low dark-current photomultiplier. Electrons spontaneously emitted by the photomultiplier photocathode contribute to the background signal of the instrument, and must, for highest sensitivity, be minimized. Fortunately the overall design of such instruments permits the photomultiplier photocathode to be of very small area, so that this particular source of background noise is not only small, but can be expected to reduce in relative importance with future improvement in photomultiplier design. Meanwhile current instruments already display very high sensitivity of detection of fluorescent signals. For example, the confocal microscope manufactured by Zeiss is claimed to display a lower detection limit for fluorescein of about ten molecules/μm^2 (Ploem, 1986). Most commercially available FITC-labelled IgG attains a fluorophore/protein molar ratio of ~4; thus the detection limit (D^*_{min}) of the Zeiss microscope is ~2–3 FITC-labelled IgG molecules/μm^2. This implies an analyte-concentration detection limit of ~2.4×10^5 molecules/ml for a two-site assay, assuming the same parameter values as used in the examples discussed above, or 2.4×10^4 molecules/ml using a 'sensing' antibody of affinity 10^{12} L/M.

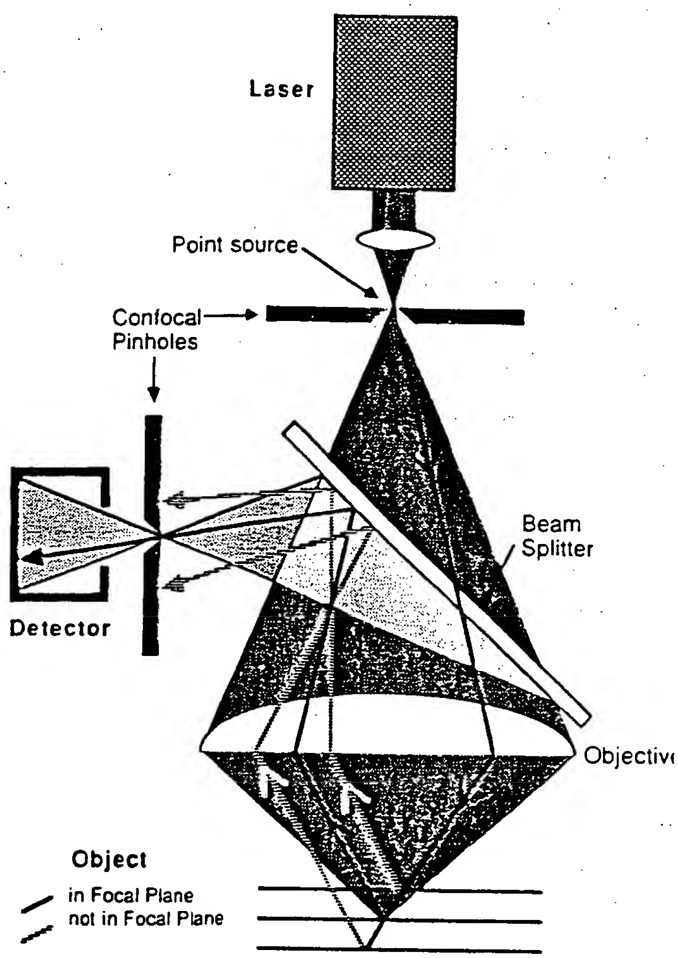


Figure 12. Principle of the confocal microscope. Illuminating light is focused at a point in the focal plane. Reflected light from this point is focused onto a detector. A complete two-dimensional image of structures within the focal plane is obtained by scanning the selected area of interest, and may be stored in a microcomputer for video display

Another comparable instrument is the Bio-Rad/Lasersharp laser scanning confocal microscope, which we are currently using in the development of 'ratiometric' multi-analyte assay methodology in accordance with the principles outlined above (see Fig. 13). The argon laser in this system possesses two excitation lines at 488 and 514 nm. It is thus particularly efficient for the excitation of blue/green emitting fluorophores such as FITC (which displays an excitation maximum at 492 nm). However, it is considerably less efficient in the excitation of red-emitting fluorophores such as Texas red (excitation maximum 596 nm). However, the ratiometric immunoassay principle permits considerable variation in detection efficiencies of the two labels relied on since, *inter alia*, the specific activities of the two labelled antibody species forming the antibody couplets can be chosen to yield optimal signal ratios in the region of unity. Thus inefficiency of the argon laser in exciting red emitting fluorophores is not necessarily a major handicap in the present context.

Though the current Lasersharp instrument relies on a conventional microscope rather than a purpose-designed optical system (and appears to

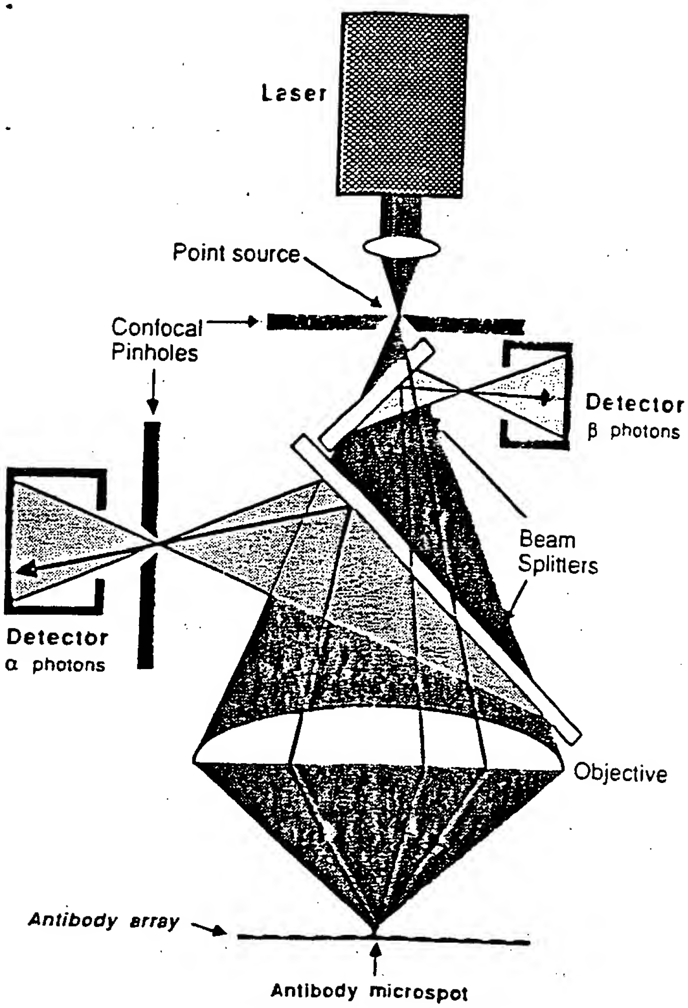


Figure 13. Dual-channel confocal fluorescence microscope permitting simultaneous measurement of the fluorescence signals from two fluorophors situated at the focal point. By scanning the antibody array, the ratio of signals from each antibody microspot may be determined

be less sensitive), it permits quantification of fluorescence signals generated from microspots of selected area. Initial studies have revealed that, under conditions that are not necessarily optimal, the instrument is capable of detecting approximately twenty-five FITC-labelled IgG molecules/μm^2, scanning an area of $\sim 50\,\mu m^2$ (Fig. 14). It must be stressed that neither of these confocal microscopes are designed specifically for routine ratiometric multi-analyte immunoassay use, and it can be anticipated that future instruments constructed specifically for this purpose are likely to prove both cheaper and more sensitive.

Other instruments. The MPM 200 Microscope Photometer manufactured by Zeiss of West Germany is anticipated to become available shortly. This photometer is claimed to be highly versatile: it can be used in transmission and reflection modes, and as a highly sensitive fluorimeter. The measuring field can be varied in shape and size for optimum adjustment to the specimen structure. More generally, the technology of sensitive light measurement is improving rapidly in response to needs in astronomy, the space program etc., such technology clearly being readily exploitable in a multi-analyte immunoassay context using light-generating labels in accordance with the broad principles presented here.

Solid antibody supports. On the basis of the theoretical considerations discussed above, it is evident that solid antibody supports for multi-analyte immunoassay use should display a capacity to adsorb a high surface density of antibody combined with low intrinsic signal-generating properties (for example, low intrinsic fluorescence), thus minimizing background. We have examined a number of materials, including polypropylene, Teflon, cellulose and nitrocellulose membranes and microtitre plates (clear polystyrene plates from Nunc; black, white and clear polystyrene plates from Dynatech with these criteria in mind. White Dynatech Microfluor microtitre plates, formulated specially for the detection of low fluorescence signals, yield high signal-to-noise ratios and have therefore been provisionally used in our developmental studies.

Surface density of antibody coating. Preliminary experiments using Microfluor plates have revealed that it is possible to coat them with antibody at a surface density of at least 5×10^4 IgG molecules/μm^2 (Fig. 15). Moreover nearly all antibody molecules so deposited appear to retain immunological activity (Fig. 16).

Verification of the 'ratiometric' imunoassay concept. Our primary intention, in initial studies, has been establishment of the basic conditions which, using a particular instrument, can be anticipated on theoretical grounds to yield high assay sensitivity. Though the setting up of individual microspot immunoassays has thus appeared to us to be of secondary importance during the initial stages of our studies, we have nevertheless

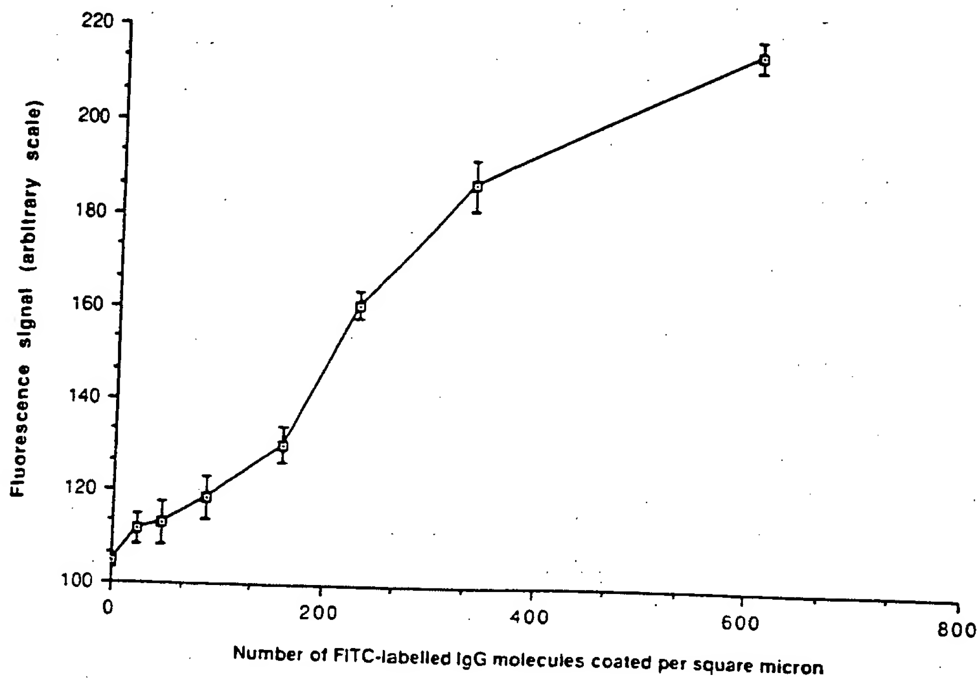


Figure 14. Fluorescence signal (arbitrary units), measured using the Bio-Rad/Lasersharp scanning confocal microscope, plotted as a function of the density of fluorescein-labelled IgG molecules (number of molecules/μm²) depositied on Dynatech Microfluor white microtitre plates

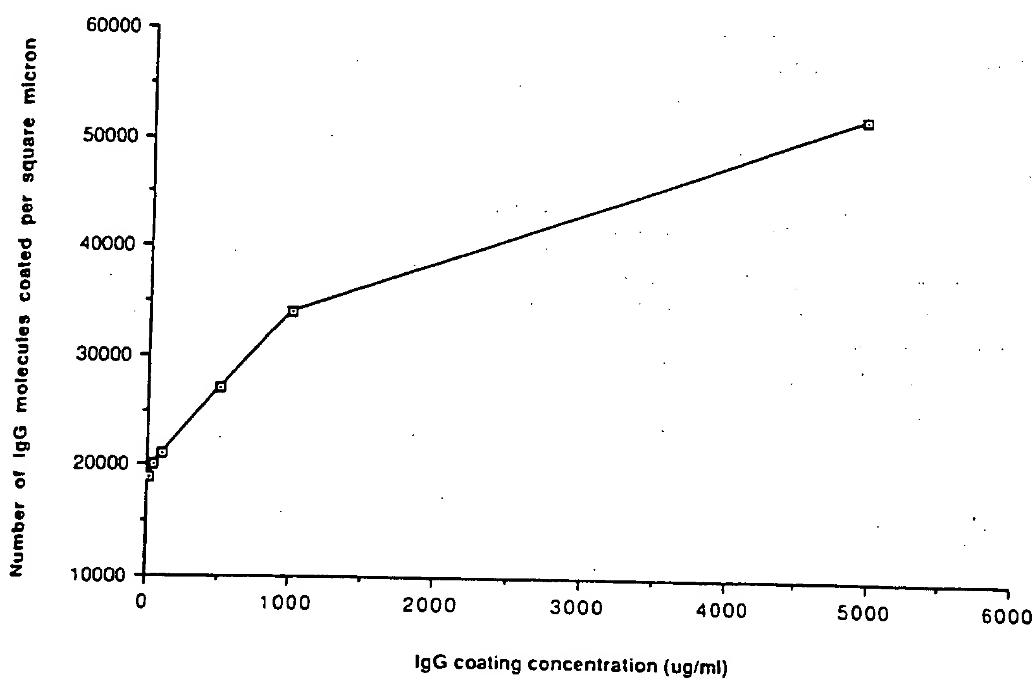


Figure 15. Surface density of IgG molecules (number of molecules/μm²) deposited on Dynatech Microfluor white plates plotted as a function of IgG concentration (μg/ml) in the coating solution

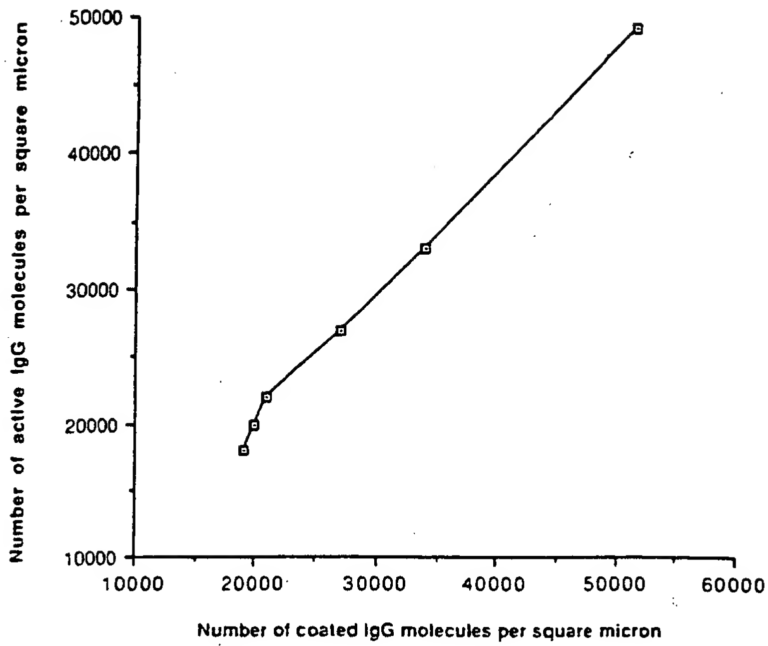


Figure 16. Surface density of immunoreactive IgG molecules (number of molecules/μm^2) plotted as a function of the total surface density of IgG (number of molecules/μm^2) on Dynatech Microfluor white microtitre plates

thought it useful to confirm the validity of our general concepts by comparing the performance of certain assays when constructed in microspot format and when conventionally designed. For example, we have compared a dual-labelled tumour necrosis factor (TNF) ratiometric assay system using Texas red and FITC-labelled antibodies with an optimized IRMA system using identical antibodies but with the second antibody ^{125}I-labelled. Although unoptimized, the ratiometric microspot assay yielded formal sensitivity values closely approaching that of the conventional, optimized, IRMA. Although verifying the general concepts underlying ratiometric microspot immunoassay methodology, further work is required to achieve the considerably greater sensitivity that theory predicts as achievable using optimized reagent concentrations and improved instrumentation.

CONCLUSION

As indicated above, differentiation of the fluorescent signals yielded by two fluorophores can be readily achieved solely on the basis of wavelength differences, and this approach has been relied on entirely in our preliminary studies. However, other physical techniques exploiting differences in decay time of two or more fluorescence emissions (using, for example, a pulsed or sinusoidally modulated laser source, and time– or phase-resolving detectors) are available, and can be expected both to further reduce background and to improve signal resolution, thus increasing assay sensitivity and precision. These considerations aside, the basic technology involved closely resembles that employed in domestic compact disk recorders and other similar data-storage devices, the obvious difference being that light emitted from each of the discrete zones forming the antibody-array is fluorescent rather than reflected, and yields chemical rather than physical information. Indeed, our preliminary studies suggest that highly sensitive immunoassays using antibody microspots of surface area approximating 50 μm^2 are achievable, implying that some 2,000,000 different immunoassays could, in principle, be accommodated on a surface area of 1 cm^2. Though non-specific binding of a multiplicity of developing antibodies would probably prohibit the use of antibody arrays of this order, it is evident that the technology is capable of encompassing analyte numbers of the kind likely to be useful in practice.

The development of multi-analyte assay systems of this kind can be anticipated to bring about

fundamental changes in medical diagnosis and many other biologically related areas. Systems capable of measuring every hormone and other endocrinologically related substance within a single small sample of blood are within technological reach, providing data which, when analysed with the aid of computer-based 'expert' pattern-recognition systems, are likely to reveal endocrine deficiences only dimly perceived using current 'single-analyte' diagnostic procedures. Such systems also provide a means to the development of a 'random access' immunoassay methodology, permitting the selection of any desired test or combination of tests from an extensive analyte menu. Clearly the accommodation of a wide range of individual immunoassays on a small immunoprobe (comparable in its overall physical dimensions with a few drops of blood) is likely to totally transform the logistics of immunodiagnostic testing, and genuinely represents, in our view, 'next generation' immunoassay methodology.

Acknowledgement

These studies are being generously supported by The Wolfson Foundation.

REFERENCES

Barnard, G. J. R., Kim, J. B. and Williams, J. L. (1985). Chemiluminescence immunoassays and immunochemiluminometric assays. In *Alternative Immunoassays*, Collins, W. P. (Ed.), John Wiley, Chichester, pp. 123–152.

Berson, S. A. and Yalow, R. S. (1973). Measurement of hormones—radioimmunoassay. In *Methods in Investigative and Diagnostic Endocrinology, 2A*, Berson, S. A. and Yalow, R. S. (Eds), North Nolland/Esevier, New York, pp. 84–135.

Dakubu, S., Ekins, R., Jackson, T. and Marshall, N. J. (1984). High sensitivity, pulsed light time-resolved fluoroimmunoassay. In *Practical Immunoassay. The State of the Art*, Butt, W. R. (Ed.), Marcel Dekker, New York, pp. 71–101.

Ekins, R. P. (1976). General principles of hormone assay. In *Hormone Assays and their Clinical Application*, Loraine, J. A. and Bell, E. T. (Eds), Churchill Livingstone, Edinburgh, pp. 1–72.

Ekins, R. P. (1978). The future development of immunoassay. In *Radioimmunoassay and Related Procedures in Medicine 1977*, IAEA, Vienna, pp.241–275.

Ekins, R. P. (1979). Radioassay methods. In *Radiopharmaceuticals II: Proceedings, 2nd International Symposium on Radiopharmaceuticals*, 19–22 March 1979, Seattle, Washington, Sorenson, J. A. (Ed), Society of Nuclear Medicine, New York, pp. 219–240.

Ekins, R. P. (1980). More sensitive immunoassays. *Nature*, **284**, 14–15.

Ekins, R. P. (1983a). The precision profile: its use in assay design, assessment and quality control. In *Immunoassays for Clinical Chemistry*, Hunter, W. M. and Corrie, J. E. T. (Eds), Churchill Livingstone, Edinburgh, pp. 76–105.

Ekins, R. P. (1983b). Measurement of analyte concentration. *British Patent* no. 8224600.

Ekins, R. (1985). Current concepts and future developments. In *Alternative Immunoassays*, Collins, W. P. (Ed.), John Wiley, Chichester, pp. 219–237.

Ekins, R. P. and Newman, B. (1970). Theoretical aspects of saturation analysis. In *Karolinska Symposia on Research Methods in Reproductive Endocrinology. 2nd Symposium: Steroid Assay by Protein Binding*, Diczfalusy, E. (Ed.), The Reproductive Endocrinology Research Unit, Karolinska sjukhuset Stockholm. pp. 11–36.

Ekins, R. P., Newman, B. and O'Riordan, J. L. H. (1968). Theoretical aspects of 'saturation' and radioimmunoassay. In *Radioisotopes in Medicine: In Vitro Studies*, Hayes, R. L., Goswitz, F. A. and Murphy, B. E. P. (Eds), Oak Ridge Symposia, USAEC, Oak Ridge, Tennessee, pp. 59–100.

Ekins, R. P., Newman, B. and O'Riordan, J. L. H. (1970a). Saturation assays. In *Statistics in Endocrinology*, McArthur, J. W. and Colton, T. (Eds), MIT Press, Cambridge, MA, pp. 345–378.

Ekins, R. P., Newman, B. and O'Riordan, J. L. H. (1970b). Competitive protein-binding assays. Discussion. In *Statistics in Endocrinology*, McArthur, J. W. and Colton, T. (Eds), MIT Press, Cambridge, MA, pp. 379–392.

Ekins, R. P., Filetti, S., Kurtz, A. B. and Dwyer, K. (1980). A simple general method for the assay of free hormones (and drugs); its application to the measurement of serum free thyroxine levels and the bearing of assay results on the 'free thyroxine' concept. *J. Endocrinol.*, **85**, 29–30.

Ezzell, C. (1986). Hybritech versus Abbott. *Nature*, **324**, 506.

Ezzell, C. (1987a). Judge confirms injunction in sandwich assay patent suit. *Nature*, **326**, 532.

Ezzell, C. (1987b). Hybritech wins court injunction over sandwich assays. *Nature*, **327**, 5.

Hemmilä, I., Dakubu, S., Mukkala, V.-M., Siiteri, H. and Lovgren, T. (1983). Europium as a label in time-resolved immunofluorometric assays. *Anal. Biochem.*, **137**, 335–343.

Jackson, T. M., Marshall, N. J. and Ekins, R. P. (1983). Optimisation of immunoradiometric (labelled antibody) assays. In *Immunoasays for Clinical Chemistry*, Hunter, W. M. and Corrie, J. E. T. (Eds), Churchill Livingstone, Edinburgh, pp. 557–575.

Kohen, F., Bayer, E. A., Wilchek, M., Barnard, G., Kim, J. B., Collins, W. P., Beheshti, I., Richardson, A. and McCapra, F. (1984). Development of luminescence-based immunoassays for haptens and for peptide hormones. In *Analytical Applications of Bioluminescence and Chemiluminescence*, Kricka, L., Stanley, P. E., Thorpe, G. H. G. and Whitehead, T. P. (Eds), Academic Press, New York, pp. 149–158.

Kohen, F., Pazzagli, M., Serio, M., DeBoever, J. and Vanderkerckhove, D. (1985). Chemiluminescence and bioluminescence immunoassay. In *Alternative Immunoassays*, Collins, W. P. (Ed). John Wiley, Chichester, pp. 103–121.

Köhler, G. and Milstein, C. (1975). Continuous culture of

fused cells secreting specific antibody. *Nature*, **256**, 495–497.
Loevinger, R. and Berman, M. (1951). Efficiency criteria in radioactive counting. *Nucleonics*, **9**, 26.
Marshall, N. J., Dakubu, S., Jackson, T. and Ekins, R. P. (1981). Pulsed-light, time-resolved, fluoroimmunoassay. In *Monoclonal Antibodies and Developments in Immunoassay*, Albertini, A. and Ekins, R. (Eds), Elsevier/North Holland, Amsterdam, pp. 101–108.
McCapra, F., Tutt, D. E. and Topping, R. M. (1977). Assay method utilizing chemiluminescence. *British Patent no.* 1, 461, 877.
McGown, L. B. and Bright, F. V. (1984). Phase-resolved fluorescence spectroscopy. *Anal. Chem.*, **56**, 1400–1417.
Miles, L. E. H. and Hales, C. N. (1968). An immunoradiometric assay of insulin. In *Protein and Polypeptide Hormones, Pt. 1*, Margoulies, M. (Ed.), Excerpta Medica, Amsterdam, pp. 61–70.
Ploem, J. S. (1986). New instrumentation for sensitive image analysis of fluorescence in cells and tissues. In *Applications of Fluorescence in the Biological Sciences*, Tayer, D. L., Waggoner, A. S., Lanni, F., Murphy, R. and Birge, R. (Eds), Alan R. Liss, New York, pp. 289–300.
Rodbard, D. and Weiss, G. H. (1973). Mathematical theory of immunometric (labelled antibody) assay. *Analyt. Biochem.*, **52**, 10–44.
Shalev, A., Greenberg, G. H. and McAlpine, P. J. (1980). Detection of attograms of antigen by a high sensitivity enzyme-linked immunosorbent assay (HS-ELISA) using a fluorogenic substrate. *J. Immunol. Methods*, **38**, 125–139.
Tait, J. F. (1970). Competitive protein-binding assays. Discussion. In *Statistics in Endocrinology*, McArthur, J. W. and Colton, T. (Eds), MIT Press, Cambridge, MA. pp. 379–392.
Weeks, I., McCapra, F., Campbell, A. K. and Woodhead, J. S. (1983). Immunoassays using chemiluminescent labelled antibodies. In *Immunoassays for Clinical Chemistry*, Hunter, W. M. and Corrie, J. E. T. (Eds), Churchill Livingstone, Edinburgh, pp. 525–530.
Weeks, I., Campbell, A. K., Woodhead, S. and McCapra, F. (1984). Immunoassays using chemiluminescent labels. In *Practical Immunoassay. The State of the Art*, Butt, W. R. (Ed.), Marcel Dekker, New York, pp. 103–116.
White, J. G., Amos, W. B. and Fordham, M. (1987). An evaluation of confocal versus conventional imaging of biological structures by fluorescence light microscopy. *J. Cell Biol.*, **105**, 41–48.
Whitehead, T. P., Thorpe, G. H., Carter, T. J., Groucutt, C. and Kricka, L. J. (1983). Enhanced luminescence procedure for sensitive determination of peroxidase-labelled conjugates in immunoassay. *Nature*, **305**, 158–159.
Wide, L. (1971). Solid phase antigen-antibody systems. In *Radioimmunoassay Methods*, Kirkham, K. E. and Hunter, W. M. (Eds), Churchill Livingstone, Edinburgh, pp. 405–418.
Woodhead, J. S., Addison, G. M., Hales, C. N. and O'Riordan, J. L. H. (1971). Discussion. In *Radioimmunoassay Methods*, Kirkham, K. E. and Hunter, W. M. (Eds), Churchill Livingstone, Edinburgh, pp. 467–488.
Yalow, R. S. and Berson, S. A. (1970a). Radioimmunoassays. In *Statistics in Endocrinology*, McArthur, J. W. and Colton, T. (Eds), MIT Press, Cambridge, MA. pp. 327–344.
Yalow, R. S. and Berson, S. A. (1970b). Cmpetitive protein-binding assays. Discussion. In *Statistics in Endocrinology*. McArthur, J. W. and Colton, T. (Eds), MIT Press, Cambridge, MA. pp. 379–392.

JOURNAL OF
BIOLUMINESCENCE AND CHEMILUMINESCENCE

Bioluminescence and Chemiluminescence: Studies and Applications in Biology and Medicine

Proceedings of the Vth International Symposium on Bioluminescence and Chemiluminescence

Editors:
M. Pazzagli, E. Cadenas, L. J. Kricka, A. Roda and P. E. Stanley

Volume 4 1989

WILEY
Chichester · New York · Brisbane · Toronto · Singapore

JBCHE7 4(1) 1–64(
ISSN 0884-3996

CLIN. CHEM. 37/11, 1955–1967 (1991)

Multianalyte Microspot Immunoassay—Microanalytical "Compact Disk" of the Future

R. P. Ekins and F. W. Chu

Throughout the 1970s, controversy centered both on immunoassay "sensitivity" per se and on the relative sensitivities of labeled antibody (Ab) and labeled analyte methods. Our theoretical studies revealed that RIA sensitivities could be surpassed only by the use of very high-specific-activity nonisotopic labels in "noncompetitive" designs, preferably with monoclonal antibodies. The time-resolved fluorescence methodology known as DELFIA—developed in collaboration with LKB/Wallac—represented the first commercial "ultrasensitive" nonisotopic technique based on these theoretical insights, the same concepts being subsequently adopted in comparable methodologies relying on the use of chemiluminescent and enzyme labels. However, high-specific-activity labels also permit the development of "multianalyte" immunoassay systems combining ultrasensitivity with the simultaneous measurement of tens, hundreds, or thousands of analytes in a small biological sample. This possibility relies on simple, albeit hitherto-unexploited, physicochemical concepts. The first is that *all* immunoassays rely on the measurement of Ab occupancy by analyte. The second is that, provided the Ab concentration used is "vanishingly small," fractional Ab occupancy is independent of both Ab concentration and sample volume. This leads to the notion of "ratiometric" immunoassay, involving measurement of the ratio of signals (e.g., fluorescent signals) emitted by two labeled Abs, the first (a "sensor" Ab) deposited as a microspot on a solid support, the second (a "developing" Ab) directed against either occupied or unoccupied binding sites of the sensor Ab. Our preliminary studies of this approach have relied on a dual-channel scanning-laser confocal microscope, permitting microspots of area 100 μm^2 or less to be analyzed, and implying that an array of 10^6 Ab-containing microspots, each directed against a different analyte, could, in principle, be accommodated on an area of 1 cm^2. Although measurement of such analyte numbers is unlikely ever to be required, the ability to analyze biological fluids for a wide spectrum of analytes is likely to transform immunodiagnostics in the next decade.

Additional Keyphrases: *ratiometric immunoassays · scanning-laser confocal microscope · fluoroimmunoassay*

Immunoassay and other protein-binding assay methods based on the use of radioisotopic labels have played a major role in medicine during the past three decades. Their utility and importance have derived primarily from the structural specificity of many reactions between binding proteins and analytes and the detectability of isotopically labeled reagents, the latter endowing such techniques with "exquisite sensitivity." Recently, however, interest has increasingly focused on nonisotopic techniques based on identical analytical principles, differing only in the nature of the marker used to label the reactant (e.g., antibody or antigen), whose distribution between reacted ("bound") and unreacted ("free") fractions constitutes the assay "response."

The basic aims underlying this interest can be broadly classed under four main headings:

- avoidance of the environmental, legal, economic, and practical disadvantages of isotopic techniques (e.g., limited shelf life of isotopically labeled reagents, problems of radioactive waste disposal, cost and complexity of radioisotope counting equipment), particularly those impeding the development of, for example, simple diagnostic kits for home or doctor's office use;
- achievement of greater assay sensitivity;
- "direct" measurement of analyte concentrations by use of transducer-based "immunosensors";
- simultaneous measurement of multiple analytes ("multianalyte assay").

In this presentation I will focus primarily on the last of these objectives, using this to set out the principles underlying our present attempts to develop a new "miniaturized" technology that will permit the simultaneous measurement of an unlimited number of analytes in a small biological sample such as a single drop of blood. However, retention (and, if possible, improvement) of the high sensitivities of conventional isotopic techniques is a basic aim not only of our own studies in this area but also of most other endeavors falling under the above headings. It is therefore appropriate to preface this paper with a discussion of the general principles underlying the attainment of high binding-assay sensitivity.

Immunoassay Sensitivity: Some Basic Concepts

Definition of Assay Sensitivity

The need to establish assay conditions yielding maximal sensitivity underlay the independent construction of mathematical theories of immunoassay design by both Yalow and Berson (*1*) and Ekins et al. (*2*) in the course of the original development of these methods in the early 1960s. Regrettably, these theoretical studies led to a prolonged controversy, arising largely from the conflicting concepts of "sensitivity" adopted by the two groups (see Figure 1). Briefly, Berson and Yalow, in their many publications relating to immunoassay design (e.g., *1*, *3*), defined sensitivity as the slope of the

Department of Molecular Endocrinology, University College and Middlesex School of Medicine, Mortimer St., London W1N 8AA, U.K.

Presented at the 23rd annual Oak Ridge Conference on Advanced Analytical Concepts for the Clinical Laboratory, St. Louis, MO, April 1991.
Received May 8, 1991; accepted August 20, 1991.

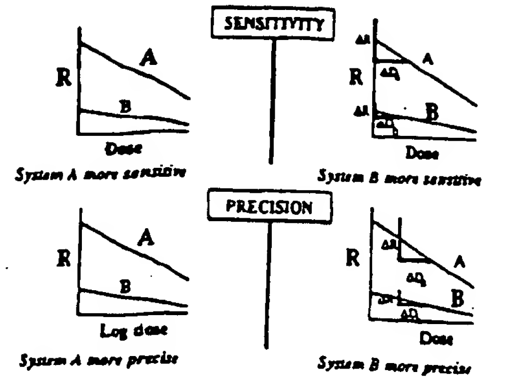


Fig. 1. The differing concepts of sensitivity and precision underlying radioimmunoassay design theories developed by (*left*) Yalow and Berson (e.g., *1, 3*) and (*right*) Ekins et al. (*2, 4*)

Yalow and Berson define assay A as more sensitive because it yields a response curve of greater slope. Ekins et al. define assay B as more sensitive because the imprecision of measurement of zero dose (σ_0) is less. Yalow and Berson likewise define an assay system as more precise if it yields a steeper response curve when data are plotted on a log dose scale

response curve relating the fraction or percentage of labeled antigen bound (b) to analyte concentration ([H]). In contrast, Ekins et al. (e.g., *2, 4*) defined sensitivity as the (im)precision of measurement of zero dose, this quantity being indicative of, and essentially equivalent to, the lower limit of detection.

The key difference between these two definitions clearly lies in the dependence of the assay detection limit on the error (imprecision) in the measurement of the response variable. By neglecting this crucial factor, the "response curve slope" definition leads to many obvious absurdities. For example, plotting conventional RIA data in terms of the response metameter B/F (i.e., the bound to free ratio) suggests that assay "sensitivity" is *increased* by increasing the antibody concentration in the system; however, the converse conclusion is reached if identical data are plotted in terms of F/B (see Figure 2). Observation of the shape and slopes of response curves without detailed error analysis thus constitutes a totally misleading guide to optimal immunoassay design. This approach has, however, characterized many of the studies conducted in the immunoassay field during the past 30 years, and has been the source of much

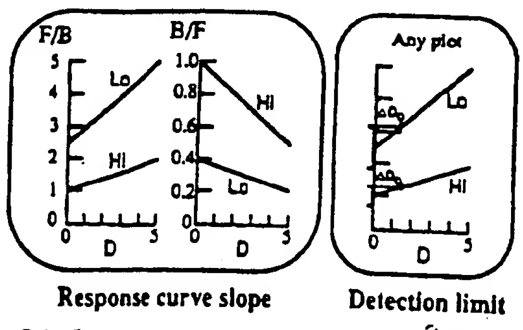


Fig. 2. Schematic representation of RIA dose–response curves observed for high and low antibody concentrations plotted in terms of (*left*) the free/bound fraction (F/B); (*center*) the bound/free fraction (B/F)

Note that the low antibody concentration yields a response curve of greater slope when the assay response is plotted in terms of F/B, but of lower slope when plotted in terms of B/F. The precision of measurement of zero dose (ΔD_0) is independent of the coordinate frame used to plot assay data (see *right*)

mythology. For example, consideration of the Law of Mass Action reveals that, when response curves corresponding to different antibody concentrations are plotted in terms of b vs [H], the maximal slope at zero dose is obtained for a concentration of 0.5/*K* (where *K* is the affinity constant), in which circumstance the zero dose response (b_0) is 33%. This conclusion led to Berson and Yalow's enunciation of the well-known dictum (which, albeit erroneous, is broadly adhered to by many immunoassay practitioners and kit manufacturers) that, to maximize RIA sensitivity, the amount of antibody to use in the system is that which binds 33% of labeled antigen in the absence of unlabeled antigen (*1, 3*).

Disagreement regarding the concept of sensitivity inevitably led to prolonged dispute regarding immunoassay design (*5*). However, although it is still common to encounter publications in the field that rely solely on the response curve slope as a measure of sensitivity, the assay detection limit is now widely accepted as the only valid indicator of this parameter, and we do not therefore intend to dwell further on this issue here. It is nevertheless relevant to an understanding of the "miniaturized" assay methodology described below to emphasize that untenable concepts of both sensitivity *and* precision underlie many of the commonly accepted rules governing current immunoassay-design practice, some of which are contravened in our own approach.

Basic Immunoassay Designs

It is likewise important in the present context to comprehend the basis of the various types of immunoassays currently in use, and the constraints on the sensitivities of which they are potentially capable. The radioimmunoassay and analogous protein-binding assay techniques originally developed for the measurement of insulin by Yalow and Berson (*6*), and of thyroxin and vitamin B_{12} by Ekins and Barakat (*7, 8*), relied on the use of a labeled analyte marker to reveal the products of the binding reactions between analyte and binder (Figure 3, left). This approach has subsequently often been portrayed as relying on "competition" between labeled and unlabeled analyte molecules for a limited number of protein-binding sites, such assays being frequently referred to as "competitive."

Subsequently, Wide et al. in Sweden (*9*), followed shortly by Miles and Hales in the U.K. (*10*), developed labeled antibody methods (Figure 3, right). These methods represented an extension of the "labeled reagent" methods (utilizing radiolabeled organic compounds such as ^{131}I-labeled *p*-iodosulfonyl chloride, [^{3}H]acetic anhydride, and other similar reagents) devised, during the early 1950s, by Keston et al. (*11*), Avivi et al. (*12*), and others for quantifying amino acids, steroid and thyroid hormones, etc. Although radiolabeled antibody methods (immunoradiometric assays; IRMAs) were originally claimed (*13*) to be more sensitive than methods based on the use of radiolabeled analyte, these claims were supported by neither rigorous theoretical analysis nor persuasive experimental evidence, and for some time remained controversial. Further doubt on their validity

analyte + antibody* ⇔ analyte : antibody* (B)
+ residual analyte
+ residual antibody* (F)

Measure "fraction bound" (B) : $[Ab]_{opt} \rightarrow \infty$
Measure "fraction free" (F) : $[Ab]_{opt} \rightarrow 0$

Fig. 3. Labeled-analyte (*left*) and labeled-antibody (*right*) assay systems compared

Labeled-analyte assay systems essentially rely on observation of an analyte "marker" to reveal the products of the reaction between analyte and antibody (although the labeled analyte is not necessarily identical to the unlabeled analyte in its binding characteristics *vis-à-vis* antibody). Note that, irrespective of which fraction of the labeled analyte is measured after the binding reaction, the optimal antibody concentration required to maximize sensitivity in such a system tends toward zero (assuming a background signal of 0). Labeled-antibody systems rely on observation of an antibody "marker" to reveal the products of the binding reaction between analyte and antibody. In this case, the optimal antibody concentration required to maximize sensitivity tends toward zero when the "free" antibody fraction is measured, but tends toward infinity when the bound fraction is determined (likewise assuming zero background)

was cast by the publication by Rodbard and Weiss in 1973 (*14*) of detailed theoretical studies demonstrating that both labeled analyte and labeled antibody methods possessed essentially equal sensitivities. (Note: These authors suggested that IRMAs might be more sensitive in the assay of small polypeptides, in which radioiodine incorporation into the antigen molecule was restricted; conversely, these assays would be less sensitive for the measurement of antigens of high molecular mass.) Nevertheless, despite the appearance of this publication, the belief that labeled antibody methods per se are intrinsically more sensitive than the corresponding labeled analyte methods gained wide acceptance among clinical chemists.

The reason for confusion on this issue is that the greater potential sensitivity of certain assay formats is not really a consequence of the labeling of antibody as opposed to analyte; indeed, the apparent antithesis between labeled-analyte and labeled-antibody methods diverts attention from the true reasons underlying the superior sensitivity of certain assay designs. Theoretical analysis (see, e.g., *4*, *15*) reveals that, assuming "perfect" separation of the products of the binding reaction (i.e., no misclassification of bound and free moieties), the optimal antibody concentration (for maximal sensitivity) in a labeled analyte immunoassay invariably tends to zero, irrespective of whether the free or bound labeled analyte fraction is measured, whereas in labeled-antibody methods the optimal antibody concentration depends on which labeled-antibody fraction is measured (see Figure 3). If the free (unreacted) antibody fraction is measured, the optimal concentration also tends to zero; conversely, if the analyte-bound fraction is measured, the concentration tends to infinity. In short, of the four basic measurement strategies available—labeled analyte, with measurement of free or bound reaction product, and labeled antibody, also with measurement of free or bound product—only one permits, in practice, the use of antibody concentrations approaching infinity.

This particular approach may, for want of a better term, be described as "noncompetitive," although it must be emphasized that such terminology involves a departure from the original meanings attached to "competitive" and "noncompetitive" when these descriptions were first used in the present context. Indeed, as discussed below, assays may be subclassified in this manner when no labeled reagent of any kind is involved.

However, the categorization of immunoassays and other binding assays as competitive or noncompetitive, depending on the binding agent concentration yielding maximal assay sensitivity, itself obscures the underlying reasons for the existence of this divergence in assay designs, and may thus be misleading. These reasons may be more readily understood if the basic principles of such assays are portrayed differently from their customary presentation.

The "Antibody Occupancy Principle" of Immunoassay

When a "sensor" antibody is introduced into an analyte-containing medium, binding sites on the antibody are occupied by analyte molecules to a fractional extent that reflects both the equilibrium constant governing the binding reaction, and the final concentration of free analyte present in the mixture. This proposition stems immediately from the Law of Mass Action, which can be written as

$$[AbAg]/[fAb] = K[fAg] \tag{1}$$

or as fractional occupancy of antibody binding sites, given by

$$[AbAg]/[Ab] = K[fAg]/(1 + K[fAg]) \tag{2}$$

where [AbAg], [Ab], [fAb], and [fAg] represent the concentrations (at equilibrium) of bound and total antibody, and free antibody and antigen (analyte), respectively, and K = equilibrium constant. The final concentration of free analyte generally depends on the concentrations of both total analyte and antibody; however, when total antibody approximates $0.05/K$ or less, free and total antigen ([Ag]) concentrations do not differ significantly, and fractional occupancy of antibody is given by

$$[AbAg]/[Ab] = K[Ag]/(1 + K[Ag]) \tag{3}$$

Assays utilizing this concept have been termed "ambient analyte immunoassays" (*16*), fractional occupancy being independent of both sample volume and antibody concentration (see below).

All immunoassays essentially depend on measurement of the "fractional occupancy" of the sensor antibody after its reaction with analyte (see Figure 4). Techniques relying on the measurement of unoccupied antibody binding sites (from which antibody occupancy is implicitly deduced by subtraction) necessitate—for attainment of maximal sensitivity—the use of sensor antibody concentrations tending to zero; these assays

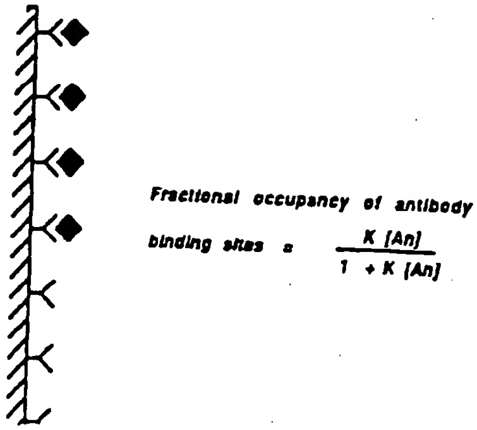


Fig. 4. The antibody binding-site occupancy principle of immunoassay
All immunoassays implicitly rely on the measurement of (fractional) binding-site occupancy by analyte

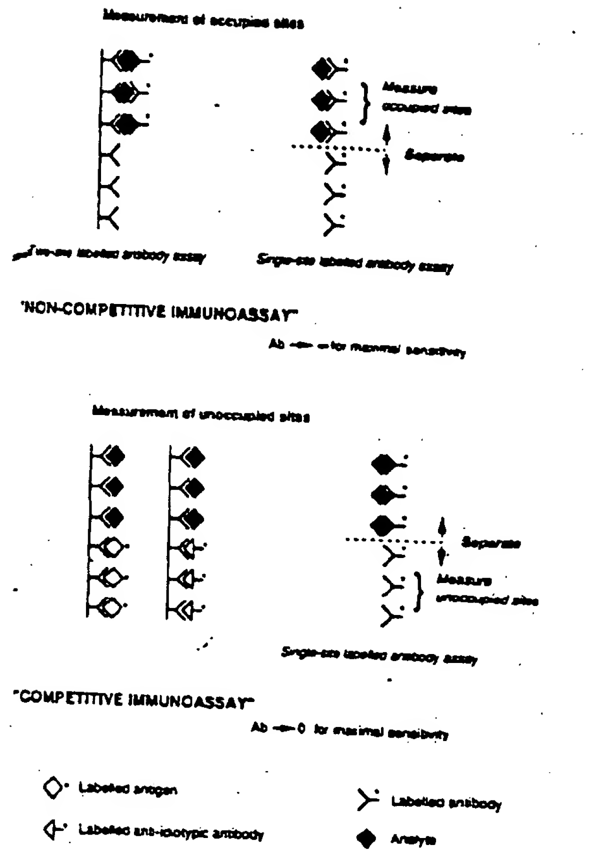


Fig. 5. Basic competitive and noncompetitive immunoassay designs
The distinction between noncompetitive and competitive immunoassays reflects the way in which antibody binding-site occupancy is observed. Labeled-antibody methods are "noncompetitive" if occupied sites of the (labeled) antibody are directly measured, but are "competitive" (*lower right*) when unoccupied sites are measured. Labeled-antigen (*lower left*) or labeled-anti-idiotypic-antibody methods (*lower center*) rely on measurement of sites unoccupied by analyte, and are therefore of "competitive" design

may therefore be categorized as "competitive." Conversely, techniques in which occupied sites are *directly* measured permit (in principle) the use of relatively high concentrations of sensor antibody and may be described as "noncompetitive." This difference in assay design simply reflects the proposition that, to minimize error in the measurement, it is generally undesirable to measure a small quantity by estimating the difference between two large quantities.

These concepts are illustrated in Figure 5, which portrays basic immunoassay formats currently in common use. Conventional RIA and other similar "labeled-analyte" techniques rely on measurement of *unoccupied* binding sites, generally by back-titration (either simultaneous or sequential) with labeled analyte, but anti-idiotypic antibody (reactive only with unoccupied sites on the sensor antibody) may be used for the same purpose. In the case of single-site labeled-antibody assays, the labeled antibody itself constitutes the sensor antibody; after reaction with analyte, this sensor antibody may be separated into occupied and unoccupied fractions through use of (e.g.) an immunosorbant (comprising antigen, antigen analog, or anti-idiotypic antibody linked to a solid support). If, after separation, the "signal" emitted by labeled antibody *bound* to analyte (i.e., the "occupied" fraction) is measured directly, the assay can be classed as "noncompetitive." Conversely, if one measures the labeled antibody *not bound* to analyte (i.e., that attached to the immunosorbant), then the assay is "competitive."

Two-site "sandwich" assays are clearly more complex because they rely on two antibodies and can be considered from two points of view. For our present purposes, the solid-phase antibody can be regarded as the "sensor" antibody, with the labeled antibody enabling the occupied sensor-antibody binding sites to be distinguished. Seen from this viewpoint, two-site assays may be classed as "noncompetitive."

These considerations emphasize that the differences in design distinguishing so-called competitive and noncompetitive methods are essentially unrelated to which component (if any) of the reaction system is labeled. Indeed, in the case of transducer-based "immunosensors," no component is labeled; nevertheless, the design of the immunosensor will differ significantly, depending on whether a measurable signal is yielded by occupied or unoccupied antibody binding sites situated on its surface. In short, the terms "competitive" and "noncompetitive" merely reflect alternative approaches to the determination of the occupancy of antibody binding sites and lead to differences in the optimal antibody concentration required to minimize the effects of random errors arising in the determination.

Competitive and noncompetitive immunoassays can be shown to differ significantly in many of their performance characteristics, including their sensitivities. In both types of assays, both the affinity constant (K) of the antibody and the specific activity of the label are important in determining sensitivity; however, in practice, the sensitivity of competitive assays is primarily limited by the affinity constant of the antibody, whereas the specific activity of the label is more important in noncompetitive systems. In both cases, the "experimental" or "manipulation" error in the measurement of the zero-dose response (R_0) [i.e., the relative error (σ_{R_0}/R_0) arising from pipetting and other operations, but not including the statistical signal measurement error per

se] is of key importance in determining "potential" assay sensitivity (i.e., the sensitivity obtained by assuming the specific activity of the label to be infinite, implying zero error in signal measurement). Thus the potential sensitivity of a competitive assay can be shown to be σ_{R_0}/KR_0, whereas that of a noncompetitive assay is given by $R_0\sigma_{R_0}/[Ab]KR_0$, where, in the latter case, R_0 is assumed to represent the labeled antibody misclassified as bound ($[bAb]_0$), commonly referred to as "nonspecifically bound" antibody. Thus $R_0/[Ab] = f$, the fraction of labeled antibody that is nonspecifically bound, and $R_0\sigma_{R_0}/[Ab]KR_0 = f\sigma_{R_0}/KR_0$. Assuming that the relative error (σ_{R_0}/R_0) in the measurement of the zero-dose response is approximately identical for both competitive and noncompetitive assays, it is evident from this simple analysis that the potential sensitivity of noncompetitive methods is greater than that of competitive methods by the factor f, i.e., by the fraction of labeled antibody that is "nonspecifically bound." For example, if the nonspecifically bound fraction is 0.01%, a noncompetitive strategy is potentially capable of a sensitivity 10 000-fold greater than that of a competitive approach, other factors being equal.

These findings are summarized in Figure 6 (left), which shows the relationships between sensitivity (expressed in terms of molecules per milliliter) and antibody affinity in an optimized competitive (labeled analyte) assay. For this analysis, we assume (*a*) the use of a label of infinite specific activity, and (*b*) the use of ^{125}I as a label, the radioactivity of the samples being counted for 1 min. Computations of the theoretically optimal reagent concentrations (on which calculations represented in Figure 6 rely) were based on the further assumptions that (*c*) the radioactivity of the antibody-bound labeled-analyte fraction was counted and (*d*) the (relative) "experimental error" component in the measurement of the bound fraction (σ_b/b) was 1%. Given these assumptions, the "potential" sensitivity attainable in such an assay is σ_b/Kb, where K is the affinity constant of the antibody. [For example, if the affinity constant is 10^{12} L/mol, and σ_b/b is 0.01 (1%), maximal assay sensitivity is 10^{-14} mol/L, or $\sim 6 \times 10^6$ molecules/mL.] The additional "signal measurement error" arising in consequence of counting radioactive samples for a finite time implies a loss of assay sensitivity, as shown by the upper curve in Figure 6 (left). However, the resulting loss in sensitivity is relatively small for antibodies of affinities $<10^{12}$ L/mol, and is negligible for antibodies with affinities $<10^{11}$ L/mol. In other words, if the assayist can accept individual sample counting times of 1–5 min, little improvement in sensitivity is gained by using alternative labels of higher specific activities than ^{125}I. However, similar considerations suggest that radioisotopic labels of much lower specific activity than ^{125}I (e.g., ^{3}H) may limit the sensitivities of the assays (such as steroid assays) in which they are used, notwithstanding the use of relatively long sample counting times.

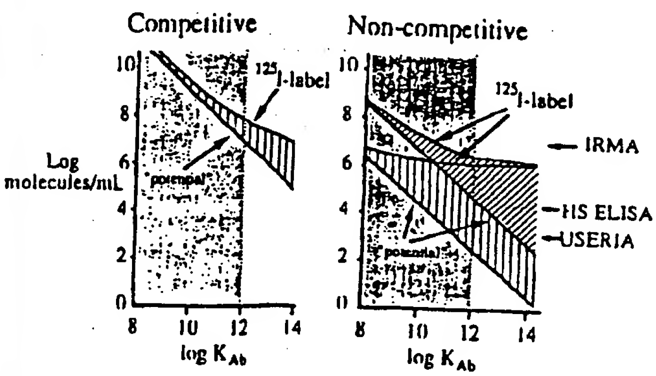


Fig. 6. Theoretically predicted sensitivities of competitive and noncompetitive immunoassay methods (represented by the SD of zero analyte measurements, expressed as molecules/mL) plotted as a function of antibody affinity (*K*)

Note: in noncompetitive sandwich assays, the antibody affinity referred to is that of the labeled antibody. In the competitive assays, calculations are based on the assumption that the experimental error (CV) incurred in the measurement of the assay response (e.g., fraction of labeled antigen bound) is 1%. The "potential sensitivity" curve assumes the use of a label of infinite specific activity, implying that the error in the measurement of the label per se is zero. The *^{125}I-label curve* indicates the loss in sensitivity arising from the statistical error incurred in counting ^{125}I disintegrations for a finite counting time. Note that, if using antibodies with an affinity $<10^{12}$ L/mol (the maximum achieved in practice), little increase in sensitivity can be achieved by using labels of higher specific activity than ^{125}I. For noncompetitive assays, the potential sensitivity curves shown relate to values of nonspecific binding of labeled antibody of 1% (*upper curves*) and 0.01% (*lower curves*), and emphasize the improvement in sensitivity potentially attainable by minimizing nonspecific binding. The corresponding *^{125}I-label curves* demonstrate the much greater loss in sensitivity (compared with that potentially attainable) when a radioisotopic marker is used, and the special advantages of nonisotopic labels of higher specific activity in noncompetitive assay designs (particularly if nonspecific binding is reduced to 0.1% or less). *Arrows* indicate assay sensitivities reported for noncompetitive immunoassays based on ^{125}I (IRMA), and enzymes relying on fluorogenic (HS-ELISA) (*28*) and radioactive (USERIA) (*29*) substrates. These conclusions underlay the original development (*19, 20*) of time-resolved fluoroimmunoassay (DELFIA), the first nonisotopic "ultra-sensitive" immunoas-

The other main conclusions stemming from such analysis are the importance of both minimizing "manipulation" errors and using antibodies of high binding affinity. For example, an increase in σ_b/b to 3% implies an approximate threefold loss in sensitivity, notwithstanding the fact that an assay reoptimized in response to the deterioration in operator skill that these numbers imply would utilize less antibody and labeled analyte, thereby partially offsetting the consequences of poor pipetting. But the most important conclusion emerging from the analysis is the near impossibility, in practice, of achieving immunoassay sensitivities better than about 10^7 molecules/mL by using a competitive approach, irrespective of the nature of the label used, if one assumes an upper limit to antibody binding affinities on the order of 10^{12} L/mol.

The results of a similar analysis of the sensitivity limitations applying to noncompetitive (two-site) assays (*15*) are illustrated in Figure 6 (right). Two sets of curves are portrayed here, corresponding to the assumptions of 1% and 0.01% nonspecific binding of labeled antibody to the capture-antibody substrate. Such analysis likewise yields important conclusions relevant to assay design, e.g., the crucial importance of reducing nonspecific binding of labeled antibody to an absolute minimum. Furthermore, if nonspecific binding is reduced to ~0.01%, just as high sensitivity is

by using an antibody of $K = 10^8$ L/mol in an optimized noncompetitive assay design as by using an antibody of $K = 10^{12}$ L/mol in a competitive method. One of the most important conclusions is that the sensitivities potentially attainable with high-affinity antibodies ($K > 10^{10}$ L/mol) are beyond the reach of radioisotopically based methods, which (because of the relatively low specific activities of isotopes such as ^{125}I) are limited in practice to sensitivities of the order of 10^6–10^7 molecules/mL or more. In short, although, under certain circumstances, noncompetitive IRMAs may be somewhat more sensitive than corresponding RIA techniques (assuming the use of the same antibody in each methodology), the potential advantages (*vis-à-vis* sensitivity) of the noncompetitive approach can be realized only by using nonisotopic labels of much higher specific activity than ^{125}I. The superiority of such labels is most apparent when they are combined with high-affinity antibodies; however, Figure 6 demonstrates that, even with use of antibodies with affinities of about 10^8–10^9 L/mol, nonisotopic labels may yield a substantial improvement in sensitivity.

These theoretical conclusions, together with the publication by Köhler and Milstein (*18*) of methods of in vitro production of monoclonal antibodies (*1*), constituted the basis of my laboratory's collaborative development (initiated around 1976) with the instrument manufacturer LKB/Wallac of the time-resolved fluorometric immunoassay methodology now known as DELFIA (*19, 20*). This methodology was the first "ultra-sensitive" nonisotopic immunoassay methodology to be developed. The same basic approach has subsequently been adopted by many other manufacturers, using a variety of high-specific activity labels (Table 1).

Against this background, let us now turn to the development of highly sensitive, miniaturized "microspot" immunoassays and multianalyte assay systems.

Antibody "Microspot" Immunoassay: Basic Concepts and Theory

Ambient Analyte Immunoassay

Particular attention has been drawn above to the specious notion that an antibody concentration approximating $0.5/K$ is required to maximize the sensitivity of conventional labeled-antigen assays. This proposition is implicitly overturned by the development of "microspot" immunoassays, which we expect to provide the basis of a new generation of binding assay methods. But before discussing this methodology in detail, another basic analytical concept must be examined.

The recognition that all immunoassays essentially rely on measurement of antibody occupancy leads to a potentially important type of assay, ambient analyte immunoassay (*16*). This name is intended to describe assay systems that, unlike conventional methods, measure the analyte *concentration* in the medium to which an antibody is exposed, being independent both of sample volume and of the amount of antibody present. The possibility of developing such assays follows from the Law of Mass Action, which leads to the following equation, representing the fractional occupancy (F) by analyte of antibody binding sites (at equilibrium):

$$F^2 - F\{(1/[\underline{Ab}]) + ([\underline{An}]/[\underline{Ab}]) + 1\} + [\underline{An}]/[\underline{Ab}] = 0 \quad (4)$$

where $[\underline{An}]$ = analyte concentration, $[\underline{Ab}]$ = antibody concentration (both in units of $1/K$).[1]

From this equation it may readily be shown that, for antibody concentrations approaching 0, $F = [\underline{An}]/(1 + [\underline{An}])$. This conclusion is illustrated in Figure 7, in which the fractional occupancy of ("monospecific" or "monoclonal") antibody binding sites in the presence of various analyte concentrations is plotted against antibody concentration. When an antibody concentration of less than (say) $0.01/K$ (the antibody preferably, but not essentially, being coupled to a solid support) is exposed to an analyte-containing medium, the resulting (fractional) occupancy of antibody binding sites solely reflects the ambient concentration of analyte[2] and is independent of the total amount of antibody in the system. (If, for example, $K = 10^{11}$ L/mol, an antibody binding-site concentration of $0.01/K$ represents 0.01×10^{-11} mol/L, or 6.02×10^7 binding sites/mL.) Analyte binding by antibody causes depletion of (unbound) analyte in the medium but, because the amount bound is small, the resulting reduction in the ambient concentration of analyte is insignificant. For example, if the concentration of binding sites of the sensor antibodies is $<0.01/K$, analyte depletion in the medium is invariably $<1\%$, and the system is therefore effectively indepen-

Table 1. Detection Limits According to Type of Label

Label	Specific activity
^{125}I	1 detectable event per second per 7.5×10^6 labeled molecules
Enzyme label	Determined by enzyme "amplification factor" and detectability of reaction product
Chemiluminescent label	1 detectable event per labeled molecule
Fluorescent label	Many detectable events per labeled molecule

[1] Expression of reagent concentrations in terms of $1/K$ units has the effect of generalizing the graphical representation of binding assay data. The terms $[\underline{Ab}]$ and $[\underline{An}]$ are underlined to indicate that this convention has been adhered to in deriving equation 4. They do not refer to molar concentrations and are not interchangeable with [Ab] and [An]. For example, if the antibody possesses an affinity (constant) for analyte of 10^{11} L/mol, a concentration of 10^{-11} mol/L (represented in units of $1/K$) is 1 (dimensionless) unit. Thus, fractional occupancy curves based on equation 4 are identical for *all* antibodies if this way of expressing antibody concentration is adopted: i.e., curves relating F to analyte concentration will be identical for systems using 10^{-11} mol/L concentrations of an antibody with an affinity of 10^{11} L/mol, 10^{-10} mol/L of an antibody with an affinity of 10^{10} L/mol, 10^{-9} mol/L of an antibody with an affinity of 10^9 L/mol, etc. (provided the analyte concentration is expressed in the same manner).

[2] The term "ambient" is used to indicate that antibody occupancy reflects the analyte concentration to which antibody binding sites are exposed, not the amount of analyte in the incubation tube; i.e., the system is independent of sample volume

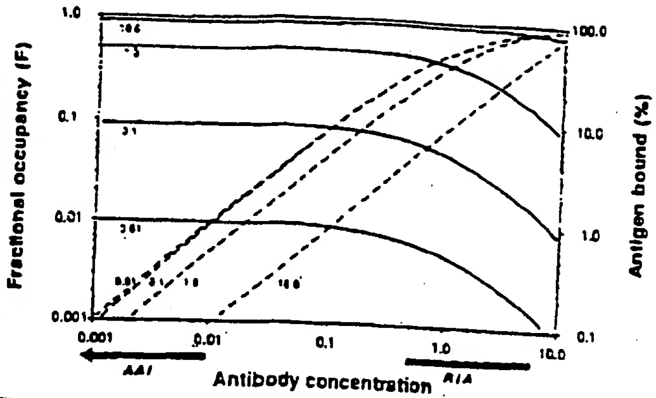


Fig. 7. Fractional antibody binding-site occupancy (F, see equation 4) plotted as a function of antibody binding-site concentration for different values of analyte (antigen) concentration (—), and the percentage binding (b) of analyte to antibody (*right-hand ordinate*; – –)

All concentrations are expressed in units of 1/*K*. Note that for antibody concentrations <0.01/*K* (approximately), the percentage binding of analyte is <1% for all analyte concentrations, and fractional binding-site occupancy is essentially unaffected by variations in antibody concentration extending over several orders of magnitude, being governed solely by antigen concentration (ambient analyte immunoassay). Note that radioimmunoassays and other "competitive" immunoassays are conventionally designed to use antibody concentrations approximating 0.5/*K*–1/*K* or more (implying binding of analyte concentrations tending to zero (b_0) >30%), in accordance with the precepts of Yalow and Berson (*1, 3*)

dent of sample volume.

These conclusions lead to two further concepts. First, the antibody may be confined to a "microspot" on a solid support, such that the total number of antibody binding sites within the microspot is $<v/K \times 10^{-5} \times N$, where v = the sample volume to which the microspot is exposed (in milliliters) and N = Avogadro's number (6×10^{23}). For example, if $v = 1$ and $K = 10^{12}$ L/mol, then the maximum number of binding sites that will cause negligible disturbance (<1%) to the ambient concentration of analyte is 6×10^{6}, this number being greater for lower-affinity antibodies. Furthermore, the perception that the ratio of occupied (or unoccupied) sites to total binding sites is solely dependent on the ambient concentration of analyte leads to the concept of a dual-label, "ratiometric," microspot immunoassay.

Dual-Label Microspot Immunoassay

After exposure of a microspot of antibody (located on a suitable probe) to an analyte-containing fluid (see Figure 8, left), the probe may be removed and exposed to a solution containing a high concentration of a "developing" antibody directed against either a second epitope (i.e., the occupied site) on the analyte molecule if the molecule is large, or against unoccupied binding sites on the antibody in the case of small analyte molecules (Figure 8, right). The fractional occupancy of the sensor antibody may thus be estimated by measuring the ratio of sensor and developing antibodies that form the dual-antibody "couplets." This can be readily achieved by labeling the sensor and the developing antibodies with different labels, e.g., a pair of radioactive, enzyme, or chemiluminescent markers (or even labels of entirely different nature). Fluorescent labels are potentially particularly useful in this context because, by the use of optical scanning techniques (Figure 9), they permit the scanning of arrays of antibody "microspots" distributed over a surface (each microspot directed against a different analyte), so that multiple analyte assays may be performed simultaneously on the same sample. Several

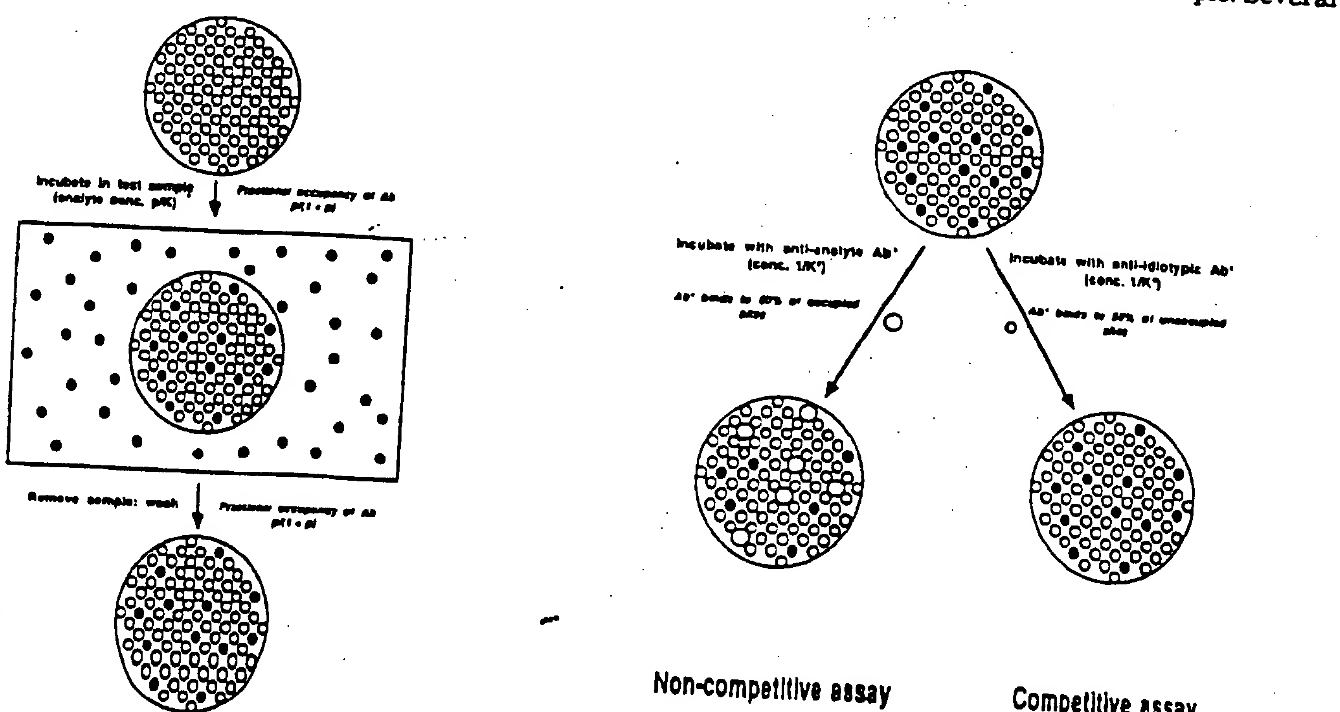


Fig. 8. Microspot immunoassay: (*left*) first incubation, with the fractional occupancy of antibody binding sites reflecting the analyte concentration to which the microspot has been exposed; (*right*) second incubation, in which the microspot is exposed to a second "developing" antibody reactive with either occupied sites (noncompetitive assay), or unoccupied sites (competitive assay)

In the second incubation, a concentration of developing antibody has been selected such that only 50% of the occupied or unoccupied sites is identified

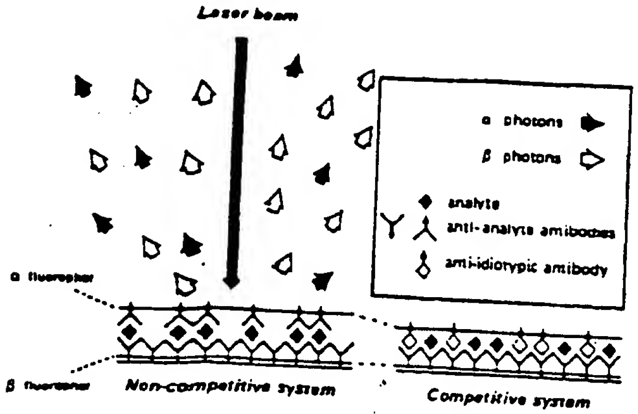


Fig. 9. Basic principle of dual-label, ambient analyte immunoassay relying on fluorescent-labeled antibodies

The ratio of α and β fluorescent photons emitted reflects the value of F (see Fig. 7) and depends solely on the analyte concentration to which the probe has been exposed. The ratio is unaffected by the amount or distribution of antibody coated (as a monomolecular layer) onto the probe surface

advantages stem from adopting a dual fluorescence measurement. For example, neither the amount nor the distribution of the sensor antibody within the detector's field of view is important, because the ratio of the emitted fluorescent signals is unaffected. Likewise, fluctuations in the intensity of the incident (exciting) light beam are apt to be of little significance. These advantages are additional to the basic benefit stemming from this approach, i.e., that the necessity of ensuring constancy of the amount of sensor antibody used in the assay system is removed.

Microspot Immunoassay Sensitivity

Because the microspot immunoassay methodology challenges concepts that have dominated immunoassay design theory in the past two to three decades, consideration of the potential sensitivity attainable by this approach is obviously of primary importance. The proposition that microspot assays may be at least as sensitive as conventional systems that rely on far larger amounts of antibody may readily be demonstrated by consideration of a model system. Let us postulate that sensor antibody molecules are attached to the surface of a solid support such that their binding sites remain exposed to the analyte, and that their affinity for the analyte is thereby unchanged. (The antibody concentration in the system—the number of binding sites on the support divided by the incubation volume—is unaffected by such attachment, and antibody occupancy by analyte at equilibrium will be identical to that occurring if the antibody is distributed uniformly throughout the incubation mixture.) Let us also suppose that the antibody molecules exist as a uniform monolayer of maximal surface density on the support and (to simplify discussion) are unlabeled. Then a change in the concentration of sensor antibody implies a corresponding change in the surface area over which the antibody is distributed. If, for example, the antibody affinity constant is 10^{11} L/mol, the total incubation volume is 1 mL, and the antibody surface density is 6000 binding sites/μm^2, then a surface area of 10^5 μm^2 (i.e., 0.1 mm^2) accommodates antibody binding sites corresponding to a concentration of $0.1/K$; an area of 0.01 mm^2 corresponds to a concentration of $0.01/K$, etc. Let us further postulate that, after exposure of the sensor antibodies to a medium containing analyte at a concentration of $0.01/K$ (i.e., 6×10^7 molecules/mL), we measure "noncompetitively" the resulting antibody occupancy (e.g., by exposure to a second, labeled, "developing" antibody directed against the analyte, forming a typical antibody sandwich). Finally, let us suppose that all occupied sites react with the developing antibody, with the latter also binding "nonspecifically" to the solid support itself at a surface density of 1 molecule/μm^2.

We may now consider the effects of a progressive reduction of the antibody-coated surface area from (e.g.) 1 mm^2 (effective antibody concentration $1/K$) through 0.1 mm^2 ($0.1/K$) to 0.01 mm^2 ($0.01/K$) and below. From equation 4, the value of F for the 1 mm^2 area is 4.98×10^{-3}. Thus at equilibrium the number of analyte and labeled antibody molecules specifically bound to the area is 2.99×10^7 (i.e., about 50% of the total analyte molecules present), whereas the number of labeled antibody molecules nonspecifically bound is 10^6. Thus, assuming the field of view of the detecting instrument is restricted to the area on which the sensor antibody is deposited (see Figure 10*a*), and (provisionally) assuming the background (or "noise") of the instrument itself to be zero (i.e., the only source of background is the non-

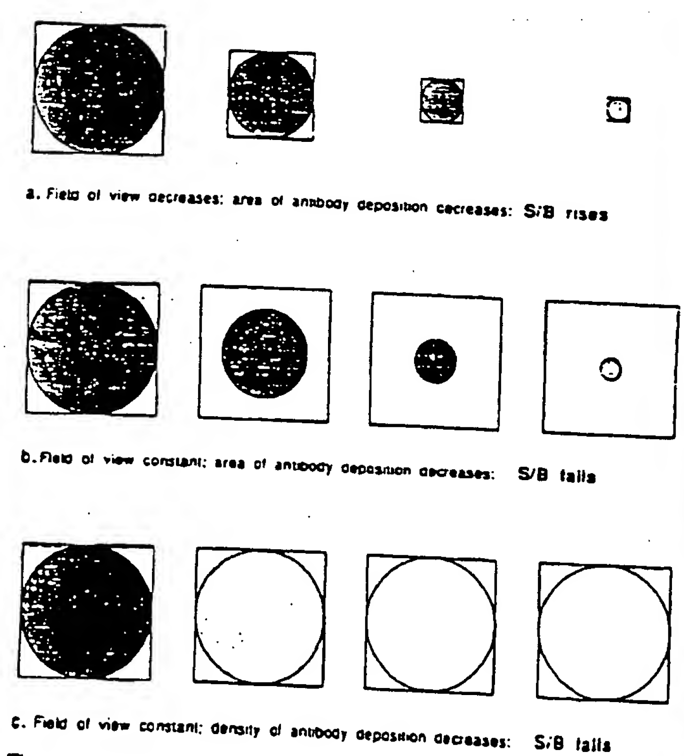


Fig. 10. "Capture" antibody (CAb) is assumed coated on circular (*shaded*) areas; the field of view of the signal-measuring instrument is represented by square (*unshaded*) areas

(*a*) Reduction of both the area of deposition of CAb and the field of view results in an increase in the signal/noise (S/B) ratio. If the CAb is reduced either by reducing the antibody coated area (*b*) or the density of antibody coating (*c*) while the field of view remains unchanged, S/B falls

specifically-bound labeled antibody within the instrument's field of view), the signal/noise ratio observed for the 1 mm^2 area is ~30. Similarly, the value of F for a 0.1 mm^2 area is 9.02×10^{-3}, the number of labeled antibody molecules specifically bound to the area is 5.41×10^6, the number nonspecifically bound is 10^5, and the signal/noise ratio is ~54. Likewise, the signal/noise ratio for a 0.01 mm^2 area can be shown to be ~59. In short, the signal/noise ratio increases as the antibody-coated surface area is decreased, approaching a maximal (plateau) value of 60 as the area coated with sensor antibody falls below 0.01 mm^2 and tends toward zero.

If, however, a reduction in the antibody-coated area were *not* accompanied by a corresponding reduction in the detecting instrument's field of view, the resulting reduction in "signal" would *not* lead to a corresponding decrease in the background generated by nonspecifically-bound developing antibody (Figure 10*b*). Therefore, although reduction in the coated area would increase the fractional occupancy of the sensor antibody, the signal/noise ratio might either remain constant or fall. In these circumstances it might be advantageous to *increase* the coated area. Similarly, if the surface density of sensor antibody were decreased (the coated area being held constant), similar conclusions would be reached (Figure 10*c*).

Likewise, if the background signal generated within the detecting instrument itself (e.g., from the photocathode of a photomultiplier tube used to detect photons emitted from the antibody-coated area) were not zero, and remained constant regardless of the instrument's field of view, then a maximum signal/noise ratio would also be attained at some optimal value of the antibody-coated area, below which the ratio would fall. Because, however, one can generally reduce the size of the detector (and hence the detector-generated background) at the same rate as the size of the signal-emitting area, there is no reason—in principle—for the signal/noise ratio to diminish as the antibody-coated area is progressively reduced toward zero. Thus if we accept the signal/noise ratio as indicative of the precision of the measurement of antibody occupancy (and hence of assay sensitivity), these considerations suggest that it is advantageous to reduce the antibody-coated surface area (and, concomitantly, the sensor-antibody concentration) toward zero, although little advantage is likely to accrue from reducing the area below 0.01 mm^2 (and thus the antibody concentration below $0.01/K$).

Were the microspot area indeed reduced to zero, both signal and noise would likewise also fall to zero (the ratio between them nevertheless remaining essentially constant), implying that no signal of any kind would, in the limit, be recorded. In practice, other statistical factors come into play when the number of individual events (e.g., photons) observed by a detecting instrument is very low, thus prohibiting a reduction of the sensor antibody concentration to zero. The point at which the reduction in the antibody-coated area causes the detectable signal to be lost sufficiently to affect the precision of the measurement of antibody occupancy depends clearly on the specific activity of the labeled antibody used to measure the occupied binding sites: the higher the specific activity, the smaller the permissible area. Thus, given labels of very high specific activity, one can envision circumstances in which, even in a "noncompetitive" system, the optimal concentration of sensor antibody may be exceedingly low. A more general conclusion is that a variety of factors, including the characteristics of the instruments used for measuring the labeled antibody (or labeled analyte), influence immunoassay design, implying, among other things, the virtual impossibility of formulating general rules regarding this. For example, reagent concentrations that are optimal for isotopically labeled reagents used with a conventional radioisotope counter (possessing a fixed background dependent on its basic construction) are likely to be entirely different when very high-specific-activity labels are used and one has the freedom to tailor the measuring instrument to samples of any size. In short, certain conclusions based on experience of RIA and IRMA techniques may prove misleading when applied to nonisotopic methodologies, and should be viewed with caution.

A more detailed theoretical consideration of (noncompetitive) microspot immunoassay sensitivity *(21)* suggests that

$$C_{\min} = D^*_{\min} \times [(6 \times 10^{20})(1 + [\mathrm{Ab}^*])]/DK[\mathrm{Ab}^*] \quad (5)$$

where D = surface density (binding sites/μm^2) of sensor antibody, K = sensor antibody affinity (L/mol), [Ab*] = concentration of labeled antibody in developing solution (expressed in units of $1/K^*$, where K^* = labeled antibody affinity), $D^*_{\min}$ = minimum detectable surface density of labeled antibody (molecules/μm^2), and $C_{\min}$ = assay detection limit (molecules/mL). For example, if [Ab*] = 1, $D = 10^5$ molecules/μm^2, $K = 10^{11}$ L/mol, and $D^*_{\min}$ = 20 molecules/μm^2, then $C_{\min} = 2.4 \times 10^6$ molecules/mL = 4×10^{-15} mol/L and the fractional occupancy of the binding sites of the sensor antibody by the minimum detectable concentration of analyte is 0.04%. Figure 11 shows the theoretical assay sensitivities attainable with use of sensor antibodies of various affinities, plotted as a function of $D^*_{\min}$.

A similar theoretical analysis of competitive microspot immunoassay indicates that potential sensitivities are essentially identical to those attainable with conventional competitive methodologies. In summary, the above considerations indicate that the attainment of high microspot assay sensitivity requires close packing of molecules of sensor antibodies within the microspot area, combined with the use of an instrument capable of accurately measuring very low surface densities of developing antibodies. They also suggest that (*a*) microspot assay sensitivities considerably higher than those obtainable by conventional isotopically based immunoassays are achievable, and (*b*) if labels of very high specific activity are available, the sensitivities yielded

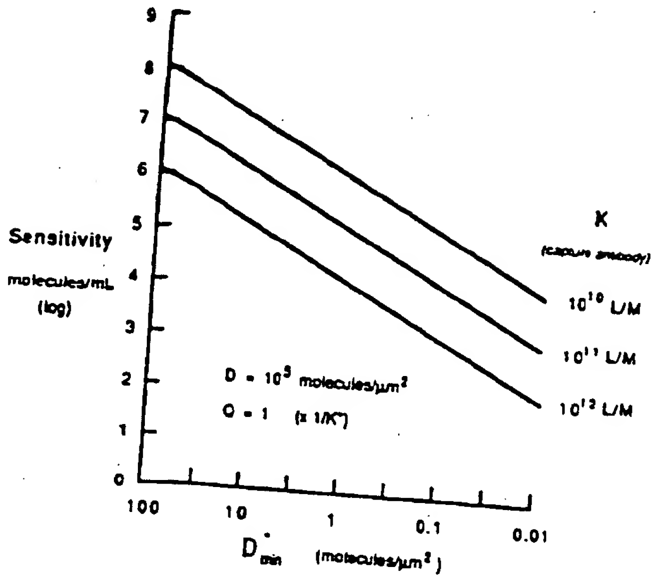


Fig. 11. Theoretically predicted sensitivity of noncompetitive microspot immunoassay plotted as a function of the minimum developing antibody density detectable within the microspot area

Postulated values of capture antibody surface density are 10^5 molecules/μm², and of developing antibody concentration are 1/K. Currently available instruments permit detection of between 10 and 1 molecules of fluorescein-labeled antibody per micrometer²

by microspot assays are unlikely to be inferior and (depending on the characteristics of the measuring instruments used) could be superior to the sensitivities achievable in macroscopic assays of conventional design.

Finally, we briefly address a further question occasionally raised in this context, i.e., the kinetic characteristics of microspot assays. Two points should be made regarding this issue. First, the smaller the microspot of sensing antibody, the lower the diffusion constraints on the velocity of the antibody/analyte binding reaction, so that at the limit (i.e., when the amount of antibody situated within the microspot area approaches zero) the kinetics of the reaction approximate those observed in a homogeneous liquid-phase system. Second, although the effective concentration of sensor antibody in the incubation medium is exceedingly low, the fractional rate at which sensor antibody binding sites within the microspot become occupied is invariably greater in this circumstance than when a relatively high concentration of antibody is used, as in conventional assays, particularly those of noncompetitive design. In other words, bearing in mind the relationship between fractional occupancy of sensor antibody and the signal/noise ratio discussed above, it is readily demonstrable that the rate at which the ratio rises is greatest when the microspot area (and the antibody contained within it) is least. Thus, given instrumentation whose field of view is restricted to the microspot area, the highest signal/noise ratio will be observed (after any selected incubation period) when the concentration of sensor antibody in the system is $<0.01/K$. In short, contrary perhaps to superficial impression, and to the generally accepted belief that short immunoassay incubation times require the use of very large amounts of antibody, the antibody microspot approach provides the basis of assays potentially more rapid than any currently available.

Microspot Immunoassay: Some Practical Considerations

Although various high-specific-activity antibody labels are potentially usable in this context, our preliminary studies have relied on the use of conventional fluorophors. The simultaneous measurement of dual fluorescences from small areas is, of course, well established, and the availability of improved instrumentation (e.g., the laser scanning confocal microscope), albeit not specifically designed for the present purpose, has been useful in demonstrating the feasibility of the microspot approach.

In laser scanning confocal fluorescence microscopes, a small area of the specimen is illuminated by a focused laser beam, the fluorescence photons emitted from this area being focused in turn onto a detector, typically a low-dark-current photomultiplier (*22, 23*). At the "confocal" point, the projection of the illumination pinhole and the back-projection of the detector pinhole coincide (Figure 12). Fluorescence photons emitted at other points thus possess a low probability of reaching the detector. Such systems contrast with conventional epifluorescence microscopes, in which the specimen is exposed to an essentially uniform flux of illumination, and yield much sharper images of fluorescent emitters situated in a defined plane of a tissue sample. Electrons spontaneously emitted by the photomultiplier photocathode contribute to the background signal of the instrument, and must—for highest microspot assay sensitivity—be minimized. Fortunately, the design of such instruments permits the photocathode to be very small in area, and this source of background can be expected

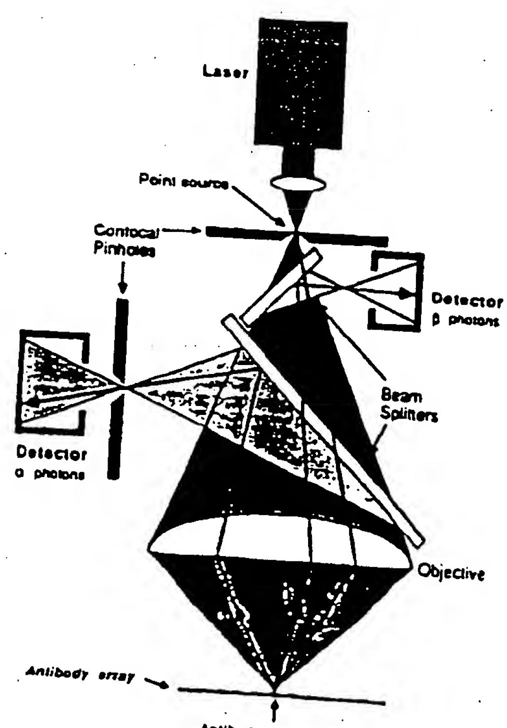


Fig. 12. Schematic diagram of a dual chan...

to diminish with future improvement in photomultiplier design. Other sources of background include fluorescence emitted by components in the optical system, which may not, in current instruments, have been constructed with background reduction as a prime consideration. Nevertheless, they detect with high sensitivity fluorescent signals. For example, one commercially available microscope is claimed to detect fluorescein at a density of 10 molecules/μm^2. Most commercially available fluorescein isothiocyanate (FITC)-labeled IgG exhibits a fluorophor/protein ratio of ~4; this implies detection limit (D^*_{min}) for antibody surface density of two or three FITC-labeled IgG molecules per micrometer^2. This, in turn, implies a theoretical sensitivity for a two-site immunoassay of ~2–3 × 10^5 analyte molecules per milliliter, assuming identical parameter values as above, or 2–3 × 10^4 molecules/mL if the sensing antibody has an affinity of 10^{12} L/mol. Clearly, sensitivity may be increased by loading more fluorophor either directly or indirectly onto the antibody.

Our preliminary studies have relied on a less sensitive microscope, albeit one possessing facilities for dual-fluorescence measurement. Its argon laser emits two excitation lines at 488 and 514 nm. It is thus particularly efficient in exciting blue/green-emitting fluorophores such as FITC (excitation maximum 492 nm), but is less efficient in exciting fluorophores such as Texas Red (excitation maximum 596 nm). However, the ratiometric assay principle permits considerable variation in detection efficiencies of the two labels because the specific activities of the labeled antibody species forming the antibody couplets can be chosen to yield signal ratios approximating unity. Inefficiency of the argon laser in exciting Texas Red is thus not a major handicap in this context. Though this instrument relies on a conventional microscope and not on an optical system designed for this purpose (and thus implicitly less sensitive), it permits quantification of fluorescence signals generated from microspots of any selected area. Initial studies have revealed that, under conditions that are not optimal, the instrument is capable of detecting ~25 FITC-labeled and (or) 150 Texas Red-labeled IgG molecules per micrometer^2, while scanning an area of ~50 μm^2.

The development of microspot immunoassays has also necessitated closer scrutiny of the mechanisms involved in the coupling of antibodies to solid supports. In the present context, these should display a capacity to adsorb (in the form of a monolayer)—or to covalently link—a high surface density of antibody combined with low intrinsic-signal-generating properties (e.g., low intrinsic fluorescence), thus minimizing background. We have examined a number of candidate materials, such as polypropylene, Teflon®, cellulose and nitrocellulose membranes, microtiter plates (clear polystyrene plates; black, white, and clear polystyrene plates), glass slides and quartz optical fibers coated with 3-(amino propyl) triethoxy silane, etc., and several alternative protocols for achieving high monolayer coating densities. These studies have exposed phenomena neither evident nor of importance when antibody binding to solid supports is examined at a macroscopic level. Provisionally, we have used white Dynatech Microfluor microtiter plates—formulated for the detection of low fluorescence signals, and yielding high signal/noise ratios and high coating densities of functional antibodies (~5 × 10^4 IgG molecules/μm^2)—for assay development, although such plates are not ideal. Indeed, deficiencies in the antibody-deposition methods used constitute the principal source of imprecision in assay results and the limitation in sensitivity that this implies. Clearly, this represents an area for further study and refinement of current coating techniques.

Notwithstanding the limitations of present instrumentation (which, among other things, does not permit the use of time-resolving techniques to distinguish two individual fluorescence signals either from each other or from background fluorescence) and the crudeness of present methods for coupling antibodies onto small areas, we have verified the theoretical concepts outlined above by comparing the performance of several assays when constructed in microspot format and when conventionally designed. Although unoptimized, ratiometric microspot assays have yielded sensitivity values closely approaching those of conventional optimized IRMA. As an example, the results of a ratiometric assay system for thyrotropin, with use of Texas Red- and FITC-labeled antibodies, are shown in Figure 13. Bearing in mind the well-known limitations of these and other "conventional" fluorophors when used as immunoassay reagent labels, such results are encouraging, although further work is clearly required to achieve the considerably greater sensitivity theoretically predicted with use of improved fluorophors, better antibody-microspotting techniques, and purpose-built (time-resolving) instrumentation.

The finding that highly sensitive immunoassays can be performed with far smaller amounts of antibody than

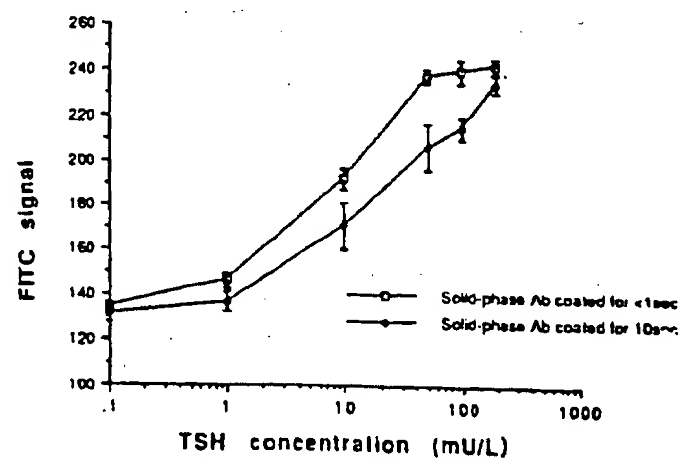


Fig. 13. Response curve in a dual-labeled microspot ratiometric assay of thyrotropin (TSH) with Texas Red-labeled solid-phase capture antibody and a developing antibody labeled with biotin/FITC–avidin

The FITC/Texas Red ratio for each microspot was measured with a scanning confocal microscope, and plotted as a function of TSH concentration in milli-int units/L

are currently used conventionally permits in turn the construction of antibody microspot arrays enabling, in principle, the simultaneous measurement of thousands of different substances in 1-mL samples. In collaboration with investigators at the Centre for Applied Microbiological Research, Porton Down, U.K., we are presently developing various techniques for the creation of such arrays. Indeed, similar technologies have recently been used for the parallel synthesis of several different polypeptides, these enabling 10 000-microspot arrays to be constructed on silica chips approximating 1 cm^2 (*24*). Although arrays of this capacity are unlikely to ever be required for conventional diagnostic purposes, we can anticipate that the ability to simultaneously measure many substances in the same sample will have revolutionary consequences in medicine and other similar areas. In addition, such techniques may ultimately permit the individual analysis of the multiple isoforms of certain "heterogeneous" analytes (e.g., the glycoprotein hormones), such molecular heterogeneity currently presenting a major obstacle to the standardization and interpretation of many immunological measurements (*25*). Moreover, although these concepts have been illustrated in an immunoassay context, they are clearly applicable to all "binding assays," including those relying on the use of DNA probes, hormone receptors, etc. For example, labeled lectins that are specific in their reactions with the sugar residues in the oligosaccharide chains of glycoprotein molecules may be used, together with specific antibodies, to impart additional "structural specificity" to sandwich assays (*26, 27*), possibly overcoming the limitations of antibodies per se in regard to differentiation of the glycosylation variants of the glycoprotein hormones.

Summary and Conclusion

Because of past confusion regarding the concepts of precision, sensitivity, accuracy, etc., several erroneous concepts have become incorporated within currently accepted rules of immunoassay design. In particular, much higher antibody concentrations are customarily used than are necessary to achieve very high assay sensitivity, provided that certain measurement strategies are adhered to. In this presentation, we have attempted to show that, in principle, the highest assay sensitivities are obtained by confining a small number of sensor antibody molecules onto a very small area in the form of a microspot and measuring their occupancy by an analyte, by using very high-specific-activity "developing" antibody probes, thereby maximizing the signal/noise ratio in the determination of sensor antibody occupancy. This observation, which contradicts currently accepted immunoassay design theory, in turn makes possible the measurement of an unlimited number of different analytes on a chip of very small surface area through the use of, e.g., laser scanning techniques closely analogous to those used in compact disk techniques of sound recording. Extensive experimental studies in this area, albeit conducted with relatively crude techniques and instrumentation not specifically designed for these purposes, and therefore not reported in detail here, have demonstrated the feasibility of the miniaturized antibody microspot approach and the validity of the general concepts on which it is based. We are therefore confident that this represents the basis of a next-generation technology that is likely to have a revolutionary impact on all fields involving the use of binding assays.

References

1. Yalow RS, Berson SA. General principles of radioimmunoassay. In: Hayes RL, Goswitz FA, Murphy BEP, eds. Radioisotopes in medicine: in vitro studies. Oak Ridge, TN: US Atomic Energy Commission, 1968:7–39.
2. Ekins RP, Newman B, O'Riordan JLH. *Ibid.*: 59–100.
3. Berson SA, Yalow RS. Measurement of hormones—radioimmunoassay. In: Berson SA, Yalow RS, eds. Methods in investigative and diagnostic endocrinology, Vol. 2A. Amsterdam: North Holland/Elsevier, 1973:84–135.
4. Ekins R, Newman B. Theoretical aspects of saturation analysis. In: Diczfalusy E, Diczfalusy A, eds. Steroid assay by protein binding. Karolinska symposia on research methods in reproductive endocrinology. Stockholm: WHO/Karolinska Sjukhuset, 1970:11–30.
5. Ekins RP. Limitations of specific activity. In: Margoulies M, ed. Protein and polypeptide hormones, Part 3 (Discussions). Amsterdam: Excerpta Medica, 1968:612–6, et seq.; Ekins RP. Concentrations of tracer and antiserum, time and temperature of incubation, volume of incubation. *Ibid*: 672–82.
6. Yalow RS, Berson SA. Immunoassay of endogenous plasma insulin in man. J Clin Invest 1960;39:1157.
7. Ekins RP. The estimation of thyroxine in human plasma by an electrophoretic technique. Clin Chim Acta 1960;5:453–9.
8. Barakat RM, Ekins RP. Assay of vitamin B_{12} in blood—a simple method. Lancet 1961;ii:25–6.
9. Wide L, Bennich H, Johansson SGO. Diagnosis of allergy by an in-vitro test for allergen antibodies. Lancet 1967;ii:1105–7.
10. Miles LEH, Hales CN. Labeled antibodies and immunological assay systems. Nature (London) 1968;219:186–9.
11. Keston AS, Udenfriend S, Cannan RK. Micro-analysis of mixtures (amino acids) in the form of isotopic derivatives. J Am Chem Soc 1946;68:1390.
12. Avivi P, Simpson SA, Tait JF, Whitehead JK. The use of ^3H and ^{14}C-labeled acetic anhydride as analytical reagents in microbiochemistry. In: Johnston JE, Faires RA, Millett RJ, eds. Radioisotope conference, London: Butterworths, 1954:313–23.
13. Miles LEH, Hales CN. An immunoradiometric assay of insulin. *Op. cit.* (ref. 5), Part 1:61–70.
14. Rodbard D, Weiss GH. Mathematical theory of immunometric (labeled antibody) assay. Anal Biochem 1973;52:10–44.
15. Jackson TM, Marshall NJ, Ekins RP. Optimisation of immunoradiometric assays. In: Hunter WM, Corrie JET, eds. Immunoassays for clinical chemistry. Edinburgh: Churchill Livingstone, 1983:557–75.
16. Ekins RP. Measurement of analyte concentration. British patent no. 8 224 600, 1983.
17. Wide L. Solid-phase antigen–antibody systems. In: Hunter WM, Kirkham KE, eds. Radioimmunoassay methods. Edinburgh: Churchill Livingstone, 1971:405–12.
18. Köhler G, Milstein C. Continuous culture of fused cells secreting specific antibody of predefined specificity. Nature (London) 1975;256:495–7.
19. Marshall NJ, Dakubu S, Jackson T, Ekins RP. Pulsed light, time resolved fluoroimmunoassay. In: Albertini A, Ekins RP, eds. Monoclonal antibodies and developments in immunoassay. Amsterdam: Elsevier/North Holland, 1981:101–8.
20. Soini E, Lövgren T. Time-resolved fluorescence of lanthanide probes and applications in biotechnology [Review]. Crit Rev Anal Chem 1987;18:105–54.
21. Ekins RP, Chu F, Biggart E. The development of microspot, multi-analyte ratiometric immunoassay using dual fluorescent-labeled antibodies. Anal Chim Acta 1990;227:73–96.
22. White JG, Amos WB, Fordham M. An evaluation of confocal versus conventional imaging of biological structures by fluores-

cence light microscopy. J Cell Biol 1987;105:41–8.
23. Ploem JS. New instrumentation for sensitive image analysis of fluorescence in cells and tissues. In: Tayer DL, Waggoner AS, Lanni F, Murphy R, Birge R, eds. Applications of fluorescence in the biological sciences. New York: Alan R Liss, 1986;289–300.
24. Fodor SPA, Read JL, Pirrung MC, et al. Light-directed, spatially addressable parallel chemical synthesis. Science 1991;251:767–73.
25. Ekins RP. Immunoassay standardization. In: Kallner A, Magid E, Albert W, eds. Improvement of comparability and compatibility of laboratory assay results in life sciences. Immunoassay standardization. Scand J Clin Lab Invest 1991;51(Suppl 205):33–46.
26. Kottgen E, Hell B, Muller C, Tauber R. Demonstration of glycosylation variants of human fibrinogen, using the new technique of glycoprotein lectin immunosorbent assay (GLIA). Biol Chem Hoppe Seyler 1988;369:1157–66.
27. Kinoshita N, Suzuki S, Matsuda Y, Taniguchi N. α-Fetoprotein antibody–lectin enzyme immunoassay to characterise sugar chains for the study of liver diseases. Clin Chim Acta 1989;179:143–52.
28. Shalev V, Greenberg GH, McAlpine PJ. Detection of attograms of antigen by a high sensitivity enzyme-linked immunosorbent assay (HS-ELISA) using a fluorogenic substrate. J Immunol Methods 1980;38:125.
29. Harris CC, Yolken RH, Kroken H, Hsu IC. Ultrasensitive enzymatic radioimmunoassay: application to detection of cholera toxin and rotavirus. Proc Natl Acad Sci USA 1979;76:5336.

Corrections

Vol 37, pp. 1447–8: In our desire for rapid publication, important errors were introduced into the following Technical Brief. The corrected version is here reproduced in its entirety, with our apologies to the authors.

Rapid Detection of 1717-1G→A Mutation in CFTR Gene by PCR-Mediated Site-Directed Mutagenesis, *Laura Cremonesi,[1] Manuela Seia,[2] Carmelina Magnani,[1] and Maurizio Ferrari[1]* ([1] Istituto Scientifico H.S. Raffaele, Lab. Centrale, Milano; [2] Istituti Clin. di Perfezionamento, Lab. di Ricerche Clin., Milano, Italy)

Until now, among the non-ΔF508 mutations identified in the cystic fibrosis transmembrane conductance regulator (CFTR) gene by the Cystic Fibrosis (CF) Genetic Analysis Consortium, the ones most frequently seen in our population sample are the 1717-1G→A mutation (13/144 or 9% of the CF chromosomes) and the G542X mutation (16/190 or 8.4% of the CF chromosomes), both revealed by dot–blot hybridization of the polymerase chain reaction (PCR) product with allele-specific oligonucleotides (ASO) probes (*1*).

In an attempt to simplify the analysis of the most frequent mutations in the CFTR gene, we converted radiolabeled ASO detection into restriction endonuclease analysis of the amplified product.

A PCR-mediated site-directed mutagenesis (*2, 3*) to detect the G542X mutation by generating a novel *BstNI* site in the wild-type sequence had already been suggested (*4*).

To detect the 1717-1G→A mutation, we designed the reverse primer (5'-CTCTGCAAACTTGGAGAGGTC-3') to contain a single-base mismatch (T→G), which could create a novel *Ava*II restriction site [G ↓ G(A/T)CC] in the amplified wild-type (WT) allele but not in the CF mutant (M) allele:

WT: WT 1717

↓

5' TAGGACA……GCAGAG 3'

3' ATCCTGG*……CGTCTC 5'

*Ava*II site

M: M 1717

↓

5' TAAGACA……GCAGAG 3'

3' ATTCTGG*……CGTCTC 5'

*, mutagenized base of reverse primer

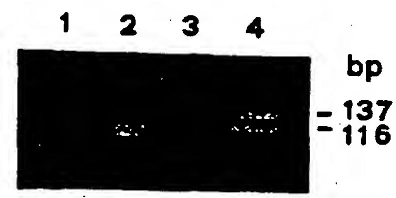


Fig. 1. Detection of the 1717-1G→A mutation by PCR

Reactions were carried out with 1 μg of genomic DNA in a total volume of 100 μL containing 10 mmol/L Tris · HCl (pH 8.3), 50 mmol/L KCl, 1.5 mmol/L $MgCl_2$, 0.1 g/L gelatin, 200 μmol/L each of the four deoxyribonucleotide triphosphates, 2.5 units of Taq polymerase (Perkin-Elmer Cetus, Norwalk, CT), and 100 pmol of each of the primers. PCR conditions were as follows: denaturation at 94 °C for 1 min, annealing at 55 °C for 30 s, and extension at 72 °C for 1 min, for a total of 30 cycles. PCR products were digested for 2 h at 37 °C with 5 U of *Ava*II and electrophoresed on 3% agarose–1% NuSieve gel for 1 h at 50 V. Bands were made visible by staining the gel with ethidium bromide. *Lane 1:* *Hae*III-digested pBR322 size marker. *Lane 2:* normal homozygote. *Lane 3:* CF patient homozygous for the 1717-1G→A mutation. *Lane 4:* heterozygote carrier for the 1717-1G→A mutation

For the forward primer, we used the one made available by the CF Genetic Analysis Consortium to amplify exon 11 of the CFTR gene: 5'-CAACTGTGGTTAAAGCAATAGTGT-3'.

Digestion by *Ava*II enzyme of the PCR product generates two fragments of 116- and 21-bp in the wild-type alleles and leaves undigested a 137-bp fragment in the mutant alleles (Figure 1).

By combined analysis for the ΔF508 mutation (*5*) (252/470 or 53.6% of the CF chromosomes), 1717-1G→A, and G542X, about 71% of mutations might be detected by nonisotopic analysis of the PCR product, thus allowing a faster and easier one-day procedure for carrier screening and prenatal testing.

References

1. Kerem B, Zielenski J, Markiewicz D, et al. Identification of mutations in regions corresponding to the two putative nucleotide (ATP)-binding folds of the cystic fibrosis gene. Proc Natl Acad Sci USA 1990;87:8447–51.
2. Haliassos A, Chomel JC, Baudis M, Kruh J, Kaplan JC, Kitzis A. Modification of enzymatically amplified DNA for the detection of point mutations. Nucleic Acids Res 1989;17:3606.
3. Friedman WE, Highsmith E Jr, Prior TW, Perry TR, Silverman LM. Cystic fibrosis deletion mutation detected by PCR-mediated site-directed mutagenesis [Tech Brief]. Clin Chem 1990;36:695–6.
4. Ng ISL, Pace R, Richard MV, et al. Methods for analysis of multiple cystic fibrosis mutations. Hum Genet (in press).
5. Ferrari M, Cremonesi L. More on detection of cystic fibrosis by polymerase chain reaction [Response to Letter]. Clin Chem 1990;36:1702–3.

11
91

clinical chemistry

In This Issue . . .

AMERICAN ASSOCIATION FOR CLINICAL CHEMISTRY
RESEARCH
SERVICE

US005432099A

United States Patent [19]
Ekins

[11] **Patent Number: 5,432,099**
[45] **Date of Patent: Jul. 11, 1995**

[54] **DETERMINATION OF AMBIENT CONCENTATION OF SEVERAL ANALYTES**

[75] Inventor: **Roger P. Ekins**, London, Great Britain

[73] Assignee: **Multilyte Limited**, United Kingdom

[21] Appl. No.: **984,264**

[22] Filed: **Dec. 1, 1992**

Related U.S. Application Data

[63] Continuation of Ser. No. 460,878, filed as PCT/GB88/00649, Aug. 5, 1988.

[30] **Foreign Application Priority Data**

Feb. 10, 1988 [GB] United Kingdom 8803000

[51] **Int. Cl.**^{6} **G01N 33/543**; G01N 33/537; G01N 33/533

[52] **U.S. Cl.** **436/518**; 435/7.1; 435/7.92; 435/973; 436/501; 436/517

[58] **Field of Search** 435/7.1, 7.92, 973; 436/501, 517, 518

[56] **References Cited**

U.S. PATENT DOCUMENTS

4,591,570 5/1986 Chang 436/518
5,096,807 3/1992 Leaback 435/6

FOREIGN PATENT DOCUMENTS

8401031 5/1984 WIPO G01N 33/54

OTHER PUBLICATIONS

Ekins et al., "Multianalyte testing", Clin Chem. 39: 369–370 (1992).
Dudley et al., "Guidelines for Immunoassay Data Processing," Clin. Chem. 31(1264–1271) (1985).

Primary Examiner—Christine M. Nucker
Assistant Examiner—M. P. Woodward
Attorney, Agent, or Firm—Dann, Dorfman, Herrell and Skillman

[57] **ABSTRACT**

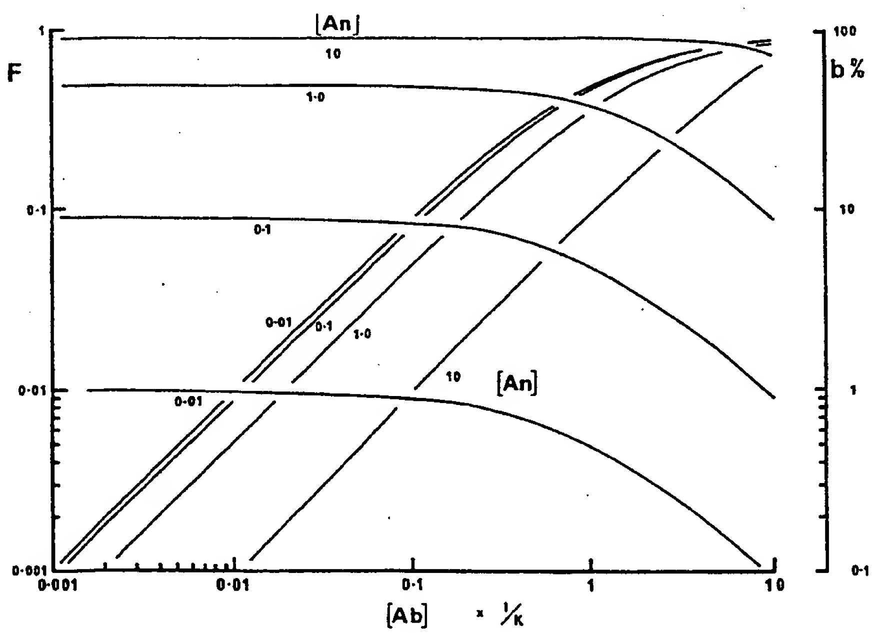
A method for determining the ambient concentrations of a plurality of analytes in a liquid sample of volume V liters, comprises

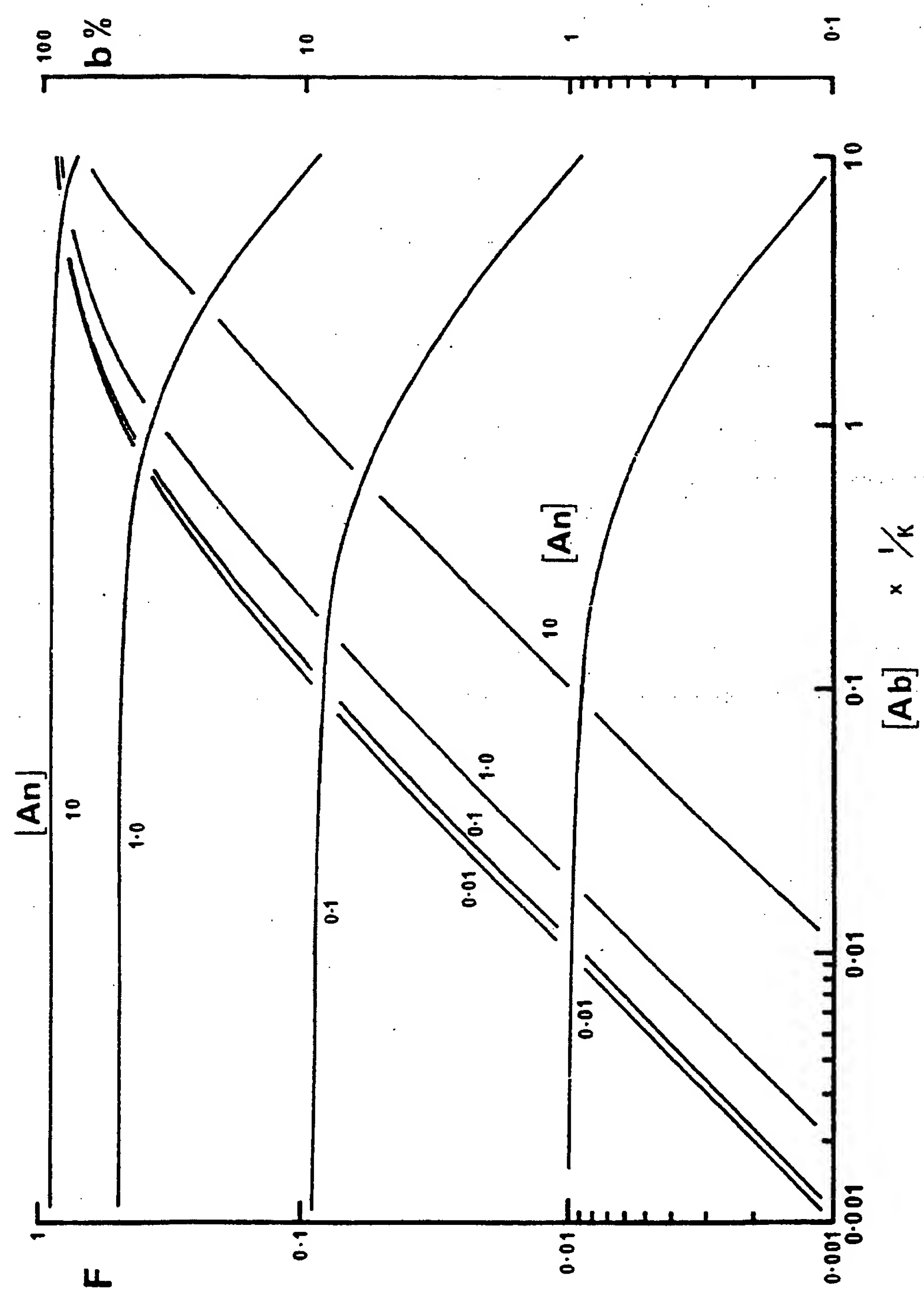
loading a plurality of different binding agents, each being capable of binding specifically and reversibly an analyte of interest onto a support means at a plurality of spaced apart locations such that not more than 0.1 V/K moles of each binding agent are present at any location, where k liters/mole is the equilibrium constant of each such binding agent;

contracting the loaded support means with the sample to be analyzed, such that each of the spaced apart locations is contacted in the same operation with the sample, the amount of sample liquid being such that only an insignificant proportion of any analyte present in the sample becomes bound to the binding agent specific for it, and

measuring a parameter representative of the fractional occupancy by the analytes of the binding agents at the spaced apart locations by a competitive or non-competitive assay technique, using a labelled site-recognition reagent for each binding agent capable of recognizing either the unfilled binding sites or the filled binding sites on the binding agent, which enables the amount of said reagent in the particular location to be measured. A device and kit for use in the method are also provided.

17 Claims, 1 Drawing Sheet




F
b %
[An]
[Ab] x 1/K
100
10
1
0·1
0·01
0·001
1·0
0·1
0·01
10

DETERMINATION OF AMBIENT CONCENTATION OF SEVERAL ANALYTES

This is a continuation of co-pending application Ser. No. 07/460,878, filed as PCT/GB88/00649, Aug. 5, 1988.

FIELD OF THE INVENTION

The present invention relates to the determination of ambient analyte concentrations in liquids, for example the determination of analytes such as hormones, proteins and other naturally occurring or artificially present substances in biological liquids such as body fluids.

BACKGROUND OF THE INVENTION

I have proposed in International Patent Application WO84/01031 to measure the concentration of an analyte in a fluid by contacting the fluid with a trace amount of a binding agent such as an antibody specific for the analyte in the sense that it reversibly binds the analyte but not other components of the fluid, determining a quantity representative of the proportional occupancy of binding sites on the binding agent and estimating from that quantity the analyte concentration. In that application I point out that, provided that the amount of binding agent is sufficiently low that its introduction into the fluid causes no significant diminution of the concentration of ambient (unbound) analyte, the fractional occupancy of the binding sites on the binding agent by the analyte is effectively independent of the absolute volume of the fluid and of the absolute amount of binding agent, i,e. independent within the limits of error usually associated with the measurement of fractional occupancy. In such circumstances, and in these circumstances only, the initial concentration [H] of analyte in the fluid is related to the fraction (Ab/Ab_o) of binding sites on the binding agent occupied by the analyte by the equation:

$$Ab/Ab_o = K_{ab}[H]/1 + K_{ab}[H]$$

where K_{ab} (hereinafter referred to as K) is the equilibrium constant for the binding of the analyte to the binding sites and is a constant for a given analyte and binding agent at any one temperature. This constant is generally known as the affinity constant, especially when the binding agent is an antibody, for example a monoclonal antibody.

The concept of using only a trace amount of binding agent is contrary to generally recommended practice in the field of immunoassay and immunometric techniques. For example, in such a well-known work as "Methods in Investigative and Diagnostic Endocrinology", ed. S. A. Berson and R. S. Yalow, 1973 at pages 111–116, it is proposed that in the performance of a competitive immunoassay maximum sensitivity of the assay is achieved if the proportion of the "tracer" analyte that is bound approximates to 50%. In order to achieve such a high degree of binding of the analyte the theory of Berson and Yalow, to this day generally accepted by other workers in the field, requires that the concentration of binding agent (or, strictly speaking, of binding sites, each molecule of binding agent conventionally having one or at most two binding sites) must be greater than or equal to the reciprocal of the equilibrium constant (K) of the binding agent for the analyte, i.e. $[ab] > 1/K$. For a sample of volume V the total amount of binding agent (or binding sites) must therefore be greater than or equal to V/K. A binding agent which is a monoclonal antibody may, for example, have an equilibrium constant (K) which is of the order of 10^{11} liters/mole for the specific antigen to which it binds. Thus, under the above generally accepted practice, a binding agent (or site) concentration of the order of 10^{-11} mole/liter or more is required for binding agents of such an equilibrium constant and, with fluid sample volumes of the order of 1 milliliter, the use of 10^{-14} or more mole of binding agent (or site) is conventionally deemed necessary. Avogadro's number is about 6×10^{23} so that 10^{-14} mole of binding site is equivalent to more than 10^9 molecules of binding agent even assuming that the binding agent possesses two binding sites per molecule. For specific binding agents of the very highest affinity K is less than 10^{13} liters/mole so that conventional practice requires more than 10^7 molecules of binding agent, whereas binding agents with lower affinity of the order of 10^8 liters/mole necessitate the use of more than 10^{12} molecules under conventional practice. In fact all immunoassay kits marketed commercially at the present time conform to these concepts and use an amount of binding site approximating to or, more frequently, considerably in excess of V/K; indeed in certain types of kit relying on the use of labelled antibodies it is conventional to use as much binding agent as possible, binding proportions of analyte greatly exceeding 50%.

Because of the binding of substantial proportions, for example 50%, of the analyte in the liquid samples under test in such systems, the fractional occupancy of the binding sites of the binding agent is not independent of the volume of the fluid sample so that for accurate quantitative assays it is necessary to control accurately the volume of the sample, keeping it constant in all tests, whether of the sample of unknown concentration or of the standard samples of known concentration used to generate the dose response curve. Furthermore, such systems also require careful control of the amount of binding agent present in the standard and control incubation tubes. These limitations of present techniques are universally recognised and accepted.

UK Patent Application 2,099,578A discloses a device for immunoassays comprising a porous solid support to which antigens, or less frequently immunoglobulins, are bound at a plurality of spaced apart locations, said device permitting a large number of qualitative or quantitative immunoassays to be performed on the same support, for example to establish an antibody profile of a sample of human blood serum. However, although the individual locations may be in the form of so-called microdots produced by supplying droplets of antigen-containing solutions or suspensions, the number of moles of antigen present at each location is apparently still envisaged as being enough to bind essentially all of the analyte (e.g. antibody) whose concentration is to be measured that is present in the liquid sample under test. This is apparent from the fact that the quantitative method used in that application (page 3, lines 21–28) involves calibration with known amounts of immunoglobulin being applied to the support; but this means that, in the samples being tested, essentially every molecule must be extracted from the sample in order for a true comparison to be made and hence that large amounts of antigen (i.e. the binding agent in this situation) are required in each microdot, greatly in excess of the total amount of analyte (i.e. antibody in this situation) present in the sample.

SUMMARY OF THE INVENTION

The present invention involves the realisation that the use of high quantities of binding agent is neither necessary for good sensitivity in immunoassays nor is it generally desirable. If, instead of being kept as large as possible, the amount of binding agent is reduced so that only an insignificant proportion of the analyte is reversibly bound to it, generally less than 10%, usually less than 5% and for optimum results only 1 or 2% or less, not only is it no longer necessary to use an accurately controlled, constant volume for all the liquid samples (standard solutions and unknown samples) in a given assay, but it is also possible to obtain reliable and sometimes even improved estimates of analyte concentration using much less than V/K moles of binding agent binding sites, say not more than 0.1 V/K and preferably less than 0.01 V/K. For a binding agent having an equilibrium constant (K) for the analyte of the order of 10^{11} liters/mole and samples of approximately 1 ml size this is approximately equivalent to not more than 10^8, preferably less than 10^7, molecules of binding agent at each location in an individual array. If the value of K is 10^{13} liters/mole the figures are 10^6 and 10^5 molecules respectively, and if K is of the order of 10^8 liters/mole they are 10^{11} and 10^{10} molecules respectively. Below 10^2 molecules of binding agent at a single location the accuracy of the measurement would become progressively less as the fractional occupancy of the binding agent sites by the analyte would be able to change only in discrete steps as individual sites become occupied or unoccupied, but in principle at least the use of as low as 10 molecules would be permissible if an estimate with an accuracy of 10% is acceptable. Practical considerations may give rise to a preference for more than 10^4 molecules.

It will be appreciated that the abovementioned GB patent application 2,099,578A, which for quantitative estimation relies on large amounts of binding agent and essentially total sequestration of all analyte, fails to recognise the advance achieved by the present invention, which instead relies on a different analytical principle requiring measurement of the fractional occupancy of the binding agent and which thus requires only a very low proportion of the total analyte molecules present to be sequestered from the sample.

Following the recognition that the use of such small amounts of binding agent is permissible, it becomes feasible to place the binding agent required for a single concentration measurement on a very small area of a solid support and hence to place in juxtaposition to one another but at spatially separate points on a single solid support a wide variety of different binding agents specific for different analytes which are or may be present simultaneously in a liquid to be analysed. Simultaneous exposure of each of the separate points to the liquid to be analysed will cause each binding agent spot to take up the analyte for which it is specific to an extent (i.e. fractional binding site occupancy) representative of the analyte concentration in the liquid, provided only that the volume of solution and the analyte concentration therein are large enough that only an insignificant fraction (generally less than 10%, usually less than 5%) of the analyte is bound to the point. The fractional binding site occupancy for each binding agent can then be determined using separate site-recognition reagents which recognise either the unfilled binding sites or filled binding sites of the different binding agents and which are labelled with markers enabling the concentration levels of the separate reagents bound to the different binding agents to be measured, for example fluorescent markers. Such measurements may be performed consecutively, for example using a laser which scans across the support, or simultaneously, for example using a photographic plate, depending on the nature of the labels. Other imaging devices such as a television camera can also be used where appropriate. Because the binding agents are spatially separate from one another it is possible to use only a small number of different marker labels or even the same marker label throughout and to scan each binding agent location separately to determine the presence and concentration of the label. By use of the invention considerably more than 3 analyses can be performed with a single exposure of the solid support with liquid to be analysed, for example 10, 20, 30, 50 or even up to 100 or several hundreds of analyses.

Overall, therefore, the present invention provides a method for determining the ambient concentrations of a plurality of analytes in a liquid sample of volume V liters, comprising:

loading a plurality of different binding agents, each being capable of reversibly binding an analyte which is or may be present in the liquid and is specific for that analyte as compared to the other components of the liquid sample, onto a support means at a plurality of spaced apart locations such that each location has not more than 0.1 V/K moles of a single binding agent, where K liters/-mole is the equilibrium constant of the binding agent for the analyte,

contacting the loaded support means with the liquid sample to be analysed such that each of the spaced apart locations is contacted in the same operation with the liquid sample, the amount of liquid used in the sample being such that only an insignificant proportion of any analyte present in the liquid sample becomes bound to the binding agent specific for it, and

measuring a parameter representative of the fractional occupancy by the analytes of the binding agents at the spaced apart locations by a competitive or non-competitive assay technique using a site-recognition reagent for each binding agent capable of recognising either the unfilled binding sites or the filled binding sites on the binding agent, said site-recognition reagent being labelled with a marker enabling the amount of said reagent in the particular location to be measured.

The invention also provides a device for use in determining the ambient concentrations of a plurality of analytes in a liquid sample of volume V liters, comprising a solid support means having located thereon at a plurality of spaced apart locations a plurality of different binding agents, each binding agent being capable of reversibly binding an analyte which is or may be present in the liquid sample and is specific for that analyte as compared to the other components of the liquid sample, each location having not more than 0.1 V/K, preferably less than 0.01 V/K, moles of a single binding agent, where K liters/mole is the equilibrium constant of that binding agent for reaction with the analyte to which it is specific.

A kit for use in the method according to the invention comprises a device according to the invention, a plurality of standard samples containing known concentrations of the analytes whose concentrations in the liquid

sample are to be measured and a set of labelled site-recognition reagents for reaction with filled or unfilled binding sites on the binding agents.

In arriving at the method of the invention, I have found that, generally speaking, for antibodies having an affinity constant K liters/mole for an antigen, the relationship between the antibody concentration and the fractional occupancy of the binding sites at any particular antigen concentration and the relationship between the antibody concentration and the percentage of antigen bound to the binding sites at any particular antigen concentration follow the same curves provided that the antibody concentrations and the antigen concentrations are each expressed in terms of fractions or multiples of 1/K.

BRIEF DESCRIPTION OF THE DRAWINGS

The principle underlying the method of the invention may be better understood by reference to the accompanying drawing which is a graph representing two sets of curves plotting the relationship between antibody concentration and the fractional occupancy of the binding sites at certain prescribed antigen concentrations and the relationship between antibody concentration and the percentage of antigen bound to the binding sites at the same prescribed antigen concentrations. Each curve relates to the antibody concentration [Ab], expressed in terms of 1/K, plotted along the x-axis. For the set of curves which remain constant or decline with increasing [Ab], the y-axis represents the fractional occupancy (F) of binding sites on the antibody by the antigen; for the second set, the y-axis represents the percentage (b%) of antigen bound to those Binding sites. The individual curves in each set represent the relationships corresponding to four different antigen concentrations [Ann] expressed in terms of K, namely 10/K, 1.0/K, O.1/K and 0.01/K. The curve show that as [Ab] falls F reaches an essentially constant level, the value of which is dependent on [An].

DETAILED DESCRIPTION

The choice of a solid support is a matter to be left to the user. Preferably the support is non-porous so that the binding agent is disposed on its surface, for example as a monolayer. Use of a porous support may cause the binding agent, depending on its molecular size, to be carried down into the pores of the support where its exposure to the analyte whose concentration is to be determined may likewise be affected by the geometry of the pores, so that a false reading may be obtained. Porous supports such as nitrocellulose paper dotted with spots of binding agent are therefore less preferred, Unlike the supports used in GB 2,099,578A, which seem to need to be porous because of the large number of molecules to be attached, the supports for use in the present invention use much smaller quantities and therefore need not be porous. The non-porous supports may, for example be of plastics material or glass, and any convenient rigid plastics material may be used, Polystyrene is a preferred plastics material, although other polyolefins or acrylic or vinyl polymers could likewise be used.

The support means may comprise microbeads, e.g. of such a plastics material, which can be coated with uniform layers of binding agent and retained in specified locations, e.g. hollows, on a support plate, Alternatively the material may be in the form of a sheet or plate which is spotted with an array of dots of binding agent, It can be advantageous for the configuration of the support means to be such that liquid samples of approximately the volume V liters are readily retained in contact with the plurality of spaced apart locations marked with the different binding agents, For example, the spaced apart locations may be arranged in a well in the support means, and a plurality of wells, each provided with the same group of different binding agents in spaced apart locations, can be linked together to form a microliter plate for use with a plurality of samples.

When the support means is to be used in conjunction with a measuring system involving light scanning, the material, e.g. plastics, for the support is desirably opaque to light, for example it may be filled with an opacifying material which may inter alia be white or black, such as carbon black, when the signals to be measured from the binding agent or the site-recognition reagent are light signals, as from fluorescent or luminescent markers. In general, reflective materials are preferred in this case to enhance light collection in the detecting instrument or photographic plate. The final choice of optimum material is governed by its ability to attach the binding agent to its surface, its absence of background signal emission and its possession of other properties tending to maximise the signal/noise ratio for the particular marker or markers attached to the binding agent situated on its surface. Very satisfactory results have been obtained in the Examples described below by the use of a white opaque polystyrene microliter plate commercially available from Dynatech under the trade name White Microfluor microliter wells.

The binding agents used may be binding agents of different specificity, that is to say agents which are specific to different analytes, or two or more of them may be binding agents of the same specificity but of different affinity, that is to say agents which are specific to the same analyte but have different equilibrium constants K for reaction with it. The latter alternative is particularly useful where the concentration of analyte to be assayed in the unknown sample can vary over considerable ranges, for example 2 or 3 orders of magnitude, as in the case of HCG measurement in urine of pregnant women, where it can vary from 0.1 to 100 or more IU/ml.

The binding agents used will preferably be antibodies, more preferably monoclonal antibodies. Monoclonal antibodies to a wide variety of ingredients of biological fluids are commercially available or may be made by known techniques. The antibodies used may display conventional affinity constants, for example from 10^8 or 10^9 liters/mole upwards, e.g. of the order of 10^{10} or 10^{11} liters/mole, but high affinity antibodies with affinity constants of 10^{12}–10^{13} liters/mole can also be used. The invention can be used with such binding agents which are not themselves labelled. However, it is also possible and frequently desirable to use labelled binding agents so that the system binding agent-/analyte/site-recognition reagent includes two different labels of the same type, e.g. fluorescent, chemiluminescent, enzyme or radioisotopic, one on the binding agent and one on the site-recognition reagent. The measuring operation then measures the ratio of the intensity of the two signals and thus eliminates the need to place the same amount of labelled binding agent on the support when measuring signals from standard samples for calibration purposes as when measuring signals from the unknown samples. Because the system depends solely on measurement of a ratio representative of binding site occupancy, there is also no need to measure the signal

from the entire spot but scanning only a portion is sufficient. Each binding agent is preferably labelled with the same label but different labels can be used.

The binding agents may be applied to the support in any of the ways known or conventionally used for coating binding agents onto supports such as tubes, for example by contacting each spaced apart location on the support with a solution of the binding agent in the form of a small drop, e.g. 0.5 microliter, on a 1 mm^2 spot, and allowing them to remain in contact for a period of time before washing the drops away. A roughly constant small fraction of the binding agent present in the drop becomes adsorbed onto the support as a result of this procedure. It is to be noted that the coating density of binding agent on the microspot does not need to be less than the coating density in conventional antibody-coated tubes; the reduction in the number of molecules on each spot may be achieved solely by reduction of the size of the spot rather than the coating density. A high coating density is generally desirable to maximise signal/noise ratios. The sizes of the spots are advantageously less than 10 mm^2 preferably less than 1 mm^2. The separation is desirably, but not necessarily, 2 or 3 times the radius of the spot, or more. These suggested geometries can nevertheless be changed as required, being subject solely to the limitations on the number of binding agent molecules in each spot, the minimum volume of the sample to which the array of spots will be exposed and the means locally available for conveniently preparing an array of spots in the manner described.

Once the binding agents have been coated onto the support it is conventional practice to wash the support, in the case of antibodies as binding agents, with a solution containing albumen or other protein to saturate all remaining non-specific adsorption sites on the support and elsewhere. To confirm that the amount of binding agent in an individual spot will be less than the maximum amount (0.1 V/K) required to conform to the principle of the present invention, the amount of binding agent present on any individual site can be checked by labelling the binding agent with a detectable marker of known specific activity (i.e. known amount of marker per unit weight of binding agent) and measuring the amount of marker present. Thus, if the use of labelled binder is not desired on the solid support used in the method of the invention the binding agent can nevertheless be labelled in a trial experiment and identical conditions to those found in that trial to give rise to correct loadings of binding agent can be used to apply unlabelled binding agent to the supports to be actually used.

The minimum size of the liquid sample (V liters) is correlated with the number of mole of binding agent (less than 0.1 V/K) so that only an insignificant proportion of the analyte present in the liquid sample becomes bound to the binding agent. This proportion is as a general rule less than 10%, usually less than 5% and desirably 1 or 2% or less, depending on the accuracy desired for the assay (greater accuracy being obtained, other things being equal, when smaller pro portions of analyte are bound) and the magnitude of other error-introducing factors present. Sample sizes of the order of one or a few ml or less, e.g. down to 100 microliters or less, are often preferred, but circumstances may arise when larger volumes are more conveniently assayed, and the geometry may be adjusted accordingly. The sample may be used at its natural concentration level or if desired it may be diluted to a known extent.

The site-recognition reagents used in the method according to the invention may themselves be antibodies, e.g. monoclonal antibodies, and may be anti-idiotypic or anti-analyte antibodies, the latter recognising occupied sites. Alternatively, for example for analytes of small molecular size such as thyroxine (T4), unoccupied sites may be recognised using either the analyte itself, appropriately labelled, or the analyte covalently coupled to another molecule—e.g. a protein molecule—which is directly or indirectly labelled. The site-recognition reagents may be labelled directly or indirectly with conventional fluorescent labels such as fluorescein, rhodamine or Texas Red or materials usable in time-resolved pulsed fluorescence such as europium and other lanthanide chelates, in a conventional manner. Other labels such as chemiluminescent, enzyme or radioisotopic labels may be used if appropriate. Each site-recognition reagent is preferably labelled with the same label but different labels can be used in different reagents. The site-recognition reagents may be specific for a single one of the binding agent/analyte spots in each group of spots or in certain circumstances, as with glycoprotein hormones such as HCG and FSH which have a common binding site, they may be cross-reacting reagents able to react with occupied binding sites in more than one of the spots.

In the assay technique the signals representative of the fractional occupancy of the binding agent in the test samples of unknown concentrations of the analytes can be calibrated by reference to dose response curves obtained from standard samples containing known concentrations of the same analytes. Such standard samples need not contain all the analytes together, provided that each of the analytes is present in some of the standard samples. Fractional occupancy may be measured by estimating occupied binding sites (as with an anti-analyte antibody) or unoccupied binding sites (as with an anti-idiotypic antibody), as one is the converse of the other. For greater accuracy it is desirable to measure the fraction which is closer to zero because a change in fractional occupancy of 0.01 is proportionately greater in this case, although for fractional occupancies in the range 25–75% either alternative is generally satisfactory.

In that embodiment of the present invention which relies on two fluorescent markers, the measurement of relative intensity of the signals from the two markers, one on the binding agent and the other on the site recognition reagent, may be carried out by a laser scanning confocal microscope such as a Bio-Rad Lasersharp MRC 500, available from Bio-Rad Laboratories Ltd., and having a dual channel detection system. This instrument relies on a laser beam to scan the dots or the like on the support to cause fluorescence of the markers and wavelength filters to distinguish and measure the amounts of fluorescence emitted. Time-resolved fluorescence methods may also be used. Interference (so-called crosstalk) between the two channels can be compensated for by standard corrections if it occurs or conventional efforts can be made to reduce it. Discrimination of the two fluorescent signals emitted by the dual-labelled spots is accomplished in the present form of this instrument, by filters capable of distinguishing the characteristic wavelength of the two fluorescent emissions; however, fluorescent substances may be distinguished by other physical characteristics such as differing fluorescence decay times, bleaching times, etc., and any of these means may be used, either alone or

in combination, to differentiate between two fluorophores and hence permit measurement of the ratio of two fluorescent labelled entities (binding agent and site-recognition reagent) present on an individual spot, using techniques well known in the fluorescence measurement field. When only one fluorescent label is present the same techniques may be used, provided that care is taken to scan the entire spot in each case and the spots contain essentially the same amount of binding agent from one assay to the next when the unknown and standard samples are used.

In the case of other labels, such as radioisotopic labels, chemiluminescent labels or enzyme labels, analogous means of distinguishing the individual signals from one or from each of a pair of such labels are also well known, For example two radioisotopes such as ^{125}I and ^{131}I may be readily distinguished on the basis of the differing energies of their respective radioactive emissions. Likewise it is possible to identify the products of two enzyme reactions, deriving from dual enzyme-labelled antibody couplets, these being e.g. of different colours, or two chemiluminescent reactions, e.g. of different chemiluminescent lifetime or wavelength of light emission, by techniques well known in the respective fields.

The invention may be used for the assaying of analytes present in biological fluids, for example human body fluids such as blood, serum, saliva or Urine. They may be used for the assaying of a wide variety of hormones, proteins, enzymes or other analytes which are either present naturally in the liquid sample or may be present artificially such as drugs, poisons or the like.

For example, the invention may be used to provide a device for quantitatively assaying a variety of hormones relating to pregnancy and reproduction, such as FSH, LH, HCG, prolactin and steroid hormones (e.g. progesterone, estradiol, testosterone and androstene-dione), or hormones of the adrenal pituitary axis, such as cortisol, ACTH and aldosterone, or thyroid-related hormones, such as T4, T3, and TSH and their binding protein TBG, or viruses such as hepatitis, AIDS or herpes virus, or bacteria, such as staphylococci, streptococci, pneumococci, gonococci and enterococci, or tumour-related peptides such as AFP or CEA, or drugs such as those banned as illicit improvers of athletes' performance, or food contaminants. In each case the binding agents used will be specific for the analytes to be assayed (as compared with others in the sample) and may be monoclonal antibodies therefor.

Further details on the methodology are to be found in my International Patent Publication W088/01058, the contents of which are incorporated herein by reference.

The invention is illustrated by the following Examples.

EXAMPLE 1

An anti-TNF (tumour necrosis factor) antibody having an affinity constant for TNF at 25° C. of about 1×10^9 liters/mole is labelled with Texas Red. A solution of the antibody at a concentration of 80 micrograms/ml is formed and 0.5 microliter aliquots of this solution are added in the form of droplets one to each well of a Dynatech Microfluor (opaque white) filled polystyrene microliter plate having 12 wells.

An anti-HCG (human chorionic gonadotropin) antibody having an affinity constant for HCG at 25° C. of about 6×10^8 liters/mole is also labelled with Texas Red. A solution of the antibody at a concentration of 80 micrograms/ml is formed and 0.5 microliter aliquots of this solution are added in the form of droplets one to each well of the same Dynatech Microfluor microliter plate.

After addition of the droplets the plate is left for a few hours in a humid atmosphere to prevent evaporation of the droplets. During this time some of the antibody molecules in the droplets become adsorbed onto the plate. Next, the wells are washed several times with a phosphate buffer and then they are filled with about 400 microliters of a 1% albumen solution and left for several hours to saturate the residual binding sites in the wells. Thereafter they are washed again with phosphate buffer.

The resulting plate has in each of its wells two spots each of area approximately 1 mm². Measurement of the amount of fluorescence shows that in each well one spot contains about 5×10^9 molecules of anti-TNF antibody and the other contains about 5×10^9 molecules of anti-HCG antibody. The wells are designed for use with liquid samples of volume 400 microliters, so that 0.1 V/K is 4×10^{-14} moles (equivalent to 2.4×10^{10} molecules) for the anti-TNF antibody and 7×10^{-14} moles (equivalent to 4×10^{10} molecules) for the anti-HCG antibody.

EXAMPLE 2

A microliter plate prepared as described in Example 1 is used in an assay for an artificially produced solution containing TNF and HCG. A test sample of the solution, amounting to about 400 microliters, is added to one of the wells and allowed to incubate for several hours. About 400 microliters of various standard solutions containing known concentrations (0.02, 0.2, 2 and 20 ng/ml) of TNF or HCG are added to other wells of the plate and also allowed to incubate for several hours. The wells are then washed several times with buffer solution.

As site-recognition reagents there are used for the TNF spots an anti-TNF antibody having an affinity constant for TNF at 25° C. of about 1×10^{10} liters/mole and for the HCG spots an anti-HCG antibody having an affinity constant for HCG at 25° C. of about 1×10^{11} liters/mole. Both antibodies are labelled with fluorescein (FITC). 400 microliter aliquots of solutions of these labelled antibodies are added to the wells and allowed to stand for a few hours. The wells are then washed with buffer.

The resulting fluorescence ratio of each spot is quantified with a Bio-Rad Lasersharp MRC 500 confocal microscope. From the standard solutions dose response curves for TNF and HCG are built up, the figures for TNF being as follows:

TNF concentration ng/ml	$\frac{\text{FITC fluorescence}}{\text{Texas Red fluorescence}}$ on TNF spot
0.02	1.1
0.2	4.6
2	7.9
20	42.5

and those for HCG being as follows:

HCG concentration ng/ml	$\frac{\text{FITC fluorescence}}{\text{Texas Red fluorescence}}$ on HCG spot
0.02	1.8
0.2	7.2

-continued

HCG concentration ng/ml	FITC fluorescence / Texas Red fluorescence on HCG spot
2	16.0
20	28.2

The artificially produced solution was found to give ratio readings of 5.9 on the TNF spot and 10.5 on the HCG spot, correlating well with the actual concentrations of TNF (0.5 ng/ml) and HCG (0.5 ng/ml) obtained from the dose response curves.

EXAMPLE 3

Using similar procedures to those outlined in Example 1 a microliter plate containing spots of labelled anti-T4 (thyroxine) antibody (affinity constant about 1×10^{11} liters/mole at 25° C), labelled anti-TSH (thyroid stimulating hormone) antibody (affinity constant about 5×10^{9} liters/mole at 25° C.) and labelled anti-T3 (triiodothyronine) antibody (affinity constant about 1×10^{11} liters/mole at 25° C.) in each of the individual wells is produced, the spots containing less than 1×10^{-12} V moles of anti-T4 antibody or less than 2×10^{-11} V moles of anti-TSH antibody or less than 1×10^{-12} V moles of anti-T3 antibody.

The developing antibody (site-recognition reagent) for the TSH assay is an anti-TSH antibody with an affinity constant for TSH of 2×10^{10} liters/mole at 25° C. This antibody is labelled with fluorescein (FITC). The site-recognition reagents for the T4 and T3 assays are T4 and T3 coupled to poly-lysine and labelled with FITC, and they recognise the unfilled sites on their respective first antibodies.

Using 400 microliter aliquots of standard solutions containing various known amounts of T4, T3 and TSH, dose response curves are obtained by methods analogous to those in Example 2, correlating fluorescence ratios with T4, T3 and TSH concentrations. The plate is used to measure T4, T3 and TSH levels in serum from human patients with good correlation with the results obtained by other methods.

EXAMPLE 4

Using similar procedures to those outlined in Example 1 a microliter plate containing spots of first labelled anti-HCG antibody (affinity constant about 6×10^{8} liters/mole at 25° C.), second labelled anti-HCG antibody (affinity constant about 1.3×10^{11} liters/mole at 25° C.) and labelled anti-FSH (follicle stimulating hormone) antibody (affinity constant about 1.3×10^{8} liters/mole at 25° C.) in each of the individual wells is produced, the spots each containing less than 0.1 V/K moles of the respective antibody. A cross-reacting (alpha subunit) monoclonal antibody 8D10 with an affinity constant of 1×10^{11} liters/mole is used as a common developing antibody for both the HCG and the FSH assays.

Using 400 microliter aliquots of standard solutions containing various known concentrations of HCG and FSH, dose response curves are obtained by methods analogous to those in Example 2, correlating fluorescence ratios with HCG and FSH concentrations, the curve obtained with the higher affinity anti-HCG antibody giving more concentration-sensitive results at the lower HCG concentrations whereas the curve from the lower affinity anti-HCG antibody is more concentration-sensitive at the higher HCG concentrations. The plate is used to measure HCG and FSH concentrations in the urine of women in pregnancy testing, giving good correlations with results obtained by other means and achieving effective concentration measurements for HCG over a concentration range of two or three orders of magnitude by correct choice of the best HCG spot and dose response curve.

Production of labelled antibodies

The labelling of the antibodies with fluorescent labels can be carried out by a well known and standard technique, see Leslie Hudson and Frank C. Hay, "Practical Immunology", Blackwell Scientific Publications (1980), pages 11-13, for example as follows:

The monoclonal antibody anti-FSH 3G3, an FSH specific (beta subunit) antibody having an affinity constant (K) of 1.3×10^{8} liters per mole, was produced in the Middlesex Hospital Medical School, and was labelled with TRITC (rhodamine isothiocyanate) or Texas Red, giving a red fluorescence.

The monoclonal antibody anti-FSH 8D10, a cross-reacting (alpha subunit) antibody having an affinity constant (K) of 1×10^{11} liters per mole, was likewise produced in the Middlesex Hospital Medical School and was labelled with FITC (fluorescein isothiocyanate), giving a yellow-green fluorescence.

The general procedure used involved ascites fluid purification (ammonium sulphate precipitation and T-gel chromatography) followed by labelling, according to the following steps:

1.a. Ammonium sulphate purification

1. Add 4.1 ml saturated ammonium sulphate solution to 5 ml anti body preparation (culture supernatant or 1:5 diluted ascites fluid) under constant stirring (45% saturation).

2. Continue stirring for 30–90 min. Centrifuge at 2500 rpm for 30 min.

3. Discard the supernatant and dissolve the precipitate in PBS (final volume 5 ml.). Repeat Steps 1 and 2, OR.

4. Add 3.6 ml saturated ammonium sulphate (40% saturation) under constant stirring. Repeat Step 2.

5. Discard the supernatant and dissolve the pellet in the desired buffer.

6. Dialyse overnight in cold against the same buffer (using fresh, boiled-in-d/w dialysis bag).

7. Determine the protein concentration either at A_{280} or by Lowry estimation.

1.b. T-gel Chromatography: (Buffer: 1M Tris-Cl, pH 7.6. Solid potassium sulphate)

1. Clear 2 ml of ascites fluid by centrifugation at 4000 rpm.

2. Add 1M Tris-Cl solution to achieve final concentration of 0.1M.

3. Add sufficient amount of solid potassium sulphate. Final concentration: =0.5M.

4. Apply the ascite fluid to the T-gel column.

5. Wash the column with 0.1M Tris-Cl buffer containing 0.5M potassium sulphate, until protein profile (at A_{280}) returns to zero.

6. Elute the absorbed protein using 0.1M Tris-Cl buffer as the eluant.

7. Pool the fractions containing antibody activity and concentrate using Amicon 30 concentrater.

8. If HPHT purification is to be carried out, use HPHT chromatography Starting buffer during Step 7.

2. Labelling of Antibodies FITC/TRITC conjugation:

1. Dialyse the purified 1 g protein into 0.25M Carbonate-bicarbonate buffer, pH 9.0 to a concentration of 20 mg/ml.

2. Add FITC/TRITC to achieve a 1:20 ratio with protein (i.e. 0.05 mg for every 1 mg of protein).

3. Mix and incubate at 4° C. for 16–18 hrs.

4. Separate the conjugated protein from unconjugated by:

a. Sephadex G-25 chromatography for FITC label, or

b. DEAE-Sephacel chromatography for TRITC-/FITC label.

Buffer system:

PBS for (a).

0.005M Phosphate, pH 8.0 and 0.18M Phosphate, pH 8.0 for (b).

$$\frac{2.87 \times \text{O.D.495 nm}}{\text{O.D.280 nm} - 0.35 \times \text{O.D.495 nm}}$$

I claim:

1. A method for determining the ambient concentrations of a plurality of analytes in a liquid sample of volume V liters, comprising:

loading a plurality of different binding agents, each being capable of reversibly binding an analyte which is or may be present in the liquid sample and is specific for said analyte as compared to the other components of the liquid sample, onto a support means at a plurality of spaced apart small spots such that each spot has a high coating density of one of said binding agents but not more than 0.1 V/K moles of binding agent are present on any spot, where K liters/mole is the affinity constant of said binding agent for said analyte;

contacting the loaded support means with the liquid sample to be analyzed, such that each of the spots is contacted in the same step with said liquid sample, the amount of liquid used in said sample being such that only an insignificant proportion of any analyte present in said liquid sample becomes bound to said binding agent specific for said analyte, and

measuring a parameter representative of the fractional occupancy by said analytes of said binding agents at the spots by a competitive or non-competitive assay technique using a site-recognition reagent for each binding agent capable of recognizing either the unfilled binding sites or the filled binding sites on said binding agent, said site-recognition reagent being labelled with a marker enabling the amount of said reagent in the particular location to be measured.

2. A method as claimed in claim 1, wherein each of said spots has a size of less than 1 mm^2.

3. A method as claimed in claim 2, wherein each of said spots contains more than 10^4 molecules of binding agent.

4. A method as claimed in claim 3, wherein each of said spots has less than 0.01 V/K moles of binding agent.

5. A method as claimed in claim 3, wherein said binding agents used have affinity constants for said analytes of from 10^8 to 10^{13} liters per mole.

6. A method as claimed in claim 3, wherein said binding agents used have affinity constants for said analytes of the order of 10^{10} or 10^{11} liters per mole.

7. A method as claimed in claim 3, wherein the volume of said liquid sample is not more than 0.1 liter.

8. A method as claimed in claim 3, wherein the volume of said liquid sample is 400 to 1000 microliters.

9. A method as claimed in claim 1, wherein said binding agents loaded onto said support means are antibodies for the analytes whose concentrations are to be determined.

10. A method as claimed in claim 1, wherein said binding agents are labelled with markers enabling the concentration levels of said binding agents to be measured.

11. A method as claimed in claim 10, wherein said binding agents and said site-recognition reagents are labelled with fluorescent markers such that at the individual spots the assay technique for measuring fractional occupancy of the binding agents measures the ratios of the signals emitted by the fluorescent markers.

12. A device for use in determining the ambient concentrations of a plurality of analytes in a liquid sample of volume V liters, comprising a solid support means having located thereon at high coating density at a plurality of spaced apart small spots a plurality of different binding agents, each binding agent being capable of reversibly binding an analyte which is or may be present in said liquid sample and is specific for said analyte as compared to the other components of said liquid sample, each spot having not more than 0.1 V/K moles of a single binding agent, where K liters/mole is the affinity constant of said single binding agent for reaction with the analyte to which it is specific.

13. A device as claimed in claim 12, wherein each of said spots has a size of less than 1 m^2.

14. A device as claimed in claim 13, wherein each of said spots contains more than 10^4 molecules of binding agent.

15. A kit for use in determining the ambient concentration of a plurality of analytes in a liquid sample of volume V liters, comprising:

a solid support means having located thereon at high coating density at a plurality of spaced apart small spots a plurality of different binding agents, each binding agent being capable of reversibly binding an analyte which is or may be present in said liquid sample and is specific for said analyte as compared to the other components of the liquid sample, each spot having not more than 0.1 V/K moles of a single binding agent, where K liters/mole is the affinity constant of said single binding agent for reaction with the analyte to which it is specific;

a plurality of standard samples containing known concentrations of the analytes whose concentrations in the liquid sample are to be measured; and

a set of labelled site-recognition reagents for reaction with filled or unfilled binding sites on said binding agents.

16. A kit as claimed in claim 15, wherein each of said spots has a size of less than 1 mm^2.

17. A kit as claimed in claim 16, wherein each of said spots contains more than 10^4 molecules of binding agent.

* * * * *

UNITED STATES PATENT AND TRADEMARK OFFICE
CERTIFICATE OF CORRECTION

PATENT NO. : 5,432,099

DATED : July 11, 1995

INVENTOR(S) : Roger P. EKINS

It is certified that error appears in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

On the title page: Item [30]

"Foreign Application Priority Data", after "Feb. 10, 1988 [GB] United Kingdom .... 8803000" insert --Aug.6,1987 [WO] .... PCT/GB87/00558 --.

Signed and Sealed this

Tenth Day of November 1998

Attest:

BRUCE LEHMAN

Attesting Officer

Commissioner of Patents and Trademarks